

Mathematical Proofs at a Crossroad?

Cristian S. Calude¹ and Solomon Marcus²

¹ University of Auckland, New Zealand
cristian@cs.auckland.ac.nz

² Romanian Academy, Romania
solomon.marcus@imar.ro.

Abstract. For more than 2000 years, from Pythagoras and Euclid to Hilbert and Bourbaki, mathematical proofs were essentially based on axiomatic-deductive reasoning. In the last decades, the increasing length and complexity of many mathematical proofs led to the expansion of some empirical, experimental, psychological and social aspects, yesterday only marginal, but now changing radically the very essence of *proof*. In this paper, we try to organize this evolution, to distinguish its different steps and aspects, and to evaluate its advantages and shortcomings. Axiomatic-deductive proofs are not *a posteriori* work, a luxury we can marginalize nor are computer-assisted proofs *bad* mathematics. There is hope for integration!

1 Introduction

From Pythagoras and Euclid to Hilbert and Bourbaki, mathematical proofs were essentially based on axiomatic-deductive reasoning. In the last decades, the increasing length and complexity of many mathematical proofs led to the expansion of some empirical, experimental, psychological and social aspects, yesterday only marginal, but now changing radically the very essence of *proof*. Computer-assisted proofs and the multiplication of the number of authors of a proof became in this way unavoidable.

In this paper, we try to organize this evolution, to distinguish its different steps and aspects and to evaluate its advantages and shortcomings. Various criticisms of this evolution, particularly, Ian Stewart's claim according to which the use of computer programs in a mathematical proof makes it as ugly "as a telephone directory" while purely axiomatic-deductive proofs are "beautiful like Tolstoy's *War and Peace*", will be discussed.

As axiomatic-deductive proofs, computer-assisted proofs may oscillate between ugliness and beauty. The elegance of a computer-program may rival the beauty of a piece of poetry, as the author of the *Art of Computer Programming* convinced us; however, this may not exclude the possibility that a computer-program assisting a proof hides a central idea or obscures the global aspect of the proof. In particular, the program assisting a proof may not be itself "proven correct", as it happened in the proof of the four-color problem, even in the latest, improved 1996 variant.

Computer-assisted proofs are to usual axiomatic-deductive proofs what (high-school) algebraic approaches are to arithmetic approaches or what analytical approaches are to direct geometric approaches. Arithmetic and intuitive geometry make children's brains more active, but algebra and analytic geometry, leading to routine and general formulas, diminish the intellectual effort and free their brains for new, more difficult problems. Obviously, each of these approaches has advantages and shortcomings, its beauty and ugliness; they are not antithetical, but complementary. Axiomatic-deductive proofs are not *a posteriori* work, a luxury we can marginalize nor are computer-assisted proofs *bad* mathematics. There is hope for integration!

2 Proofs in General

Proofs are used in everyday life and they may have nothing to do with mathematics. There is a whole field of research, at the intersection of logic, linguistics, law, psychology, sociology, literary theory etc., concerning the way people argue: argumentation theory. Sometimes, this is a subject taught to 15 or 16 year-old students.

In the "Oxford American Dictionary" [16] we read:

Proof: 1. a fact or thing that shows or helps to show that something is true or exists; 2. a demonstration of the truth of something, "in proof of my statement"; 3. the process of testing whether something is true or good or valid, "put it to the proof". *To prove*: to give or be proof of; to establish the validity of; to be found to be, "it proved to be a good theory"; to test or stay out. *To argue*: 1. to express disagreement, to exchange angry words; 2. to give reasons for or against something, to debate; 3. to persuade by talking, "argued him into going"; 4. to indicate, "their style of living argues that they are well off". *Argument*: 1. a discussion involving disagreement, a quarrel; 2. a reason put forward; 3. a theme or chain of reasoning.

In all these statements, nothing is said about the means used "to show or help to show that something is true or exists", about the means used "in the process of testing whether something is true or good or valid". In argumentation theory, various ways to argue are discussed, deductive reasoning being only one of them. The literature in this respect goes from classical rhetorics to recent developments such as [28]. People argue by all means. We use suggestions, impressions, emotions, logic, gestures, mimicry, etc.

What is the relation between proof in general and proof in mathematics? It seems that the longer a mathematical proof is, the higher the possibility to contain elements usually belonging to non-mathematical proofs. We have in view emotional, affective, intuitive, social elements related to fatigue, memory gaps, etc. Long proofs are not necessarily computational; the proof of Fermat's theorem and the proof of Bieberbach's conjecture did not use computer programs, but they paid a price for their long lengths.

3 From Proofs to Mathematical Proofs

Why did mathematical proofs, beginning with Thales, Pythagoras and Euclid, till recently, use only deductive reasoning? First of all, deduction, syllogistic reasoning is the most visible aspect of a mathematical proof, but not the only one. Observation, intuition, experiment, visual representations, induction, analogy and examples have their role; some of them belong to the preliminary steps, whose presence is not made explicit, but without which proofs cannot be conceived. As a matter of fact, neither deduction, nor experiment could be completely absent in a proof, be it the way it was conceived in Babylonian mathematics, predominantly empirical, or in Greek mathematics, predominantly logical. The problem is one of *proportion*. In the 1970s, for the first time in the history of mathematics, empirical-experimental tools, under the form of some computer programs, have penetrated massively in mathematics and led to a solution of the four-color problem (4CP), a solution which is still an object of debate and controversy, see Appel and Haken [1], Tymoczko [39], Swart [37], Marcus [27], and A. Calude [10].³

Clearly, any proof, be it mathematical or not, is a very heterogeneous process, where different ingredients are involved in various degrees. The increasing role of empirical-experimental factors may recall the Babylonian mathematics, with the significant difference that the deductive component, today impressive, was then very poor. But what is the difference between ‘proof’ and ‘mathematical proof’? The difficulty of this question is related to the fact that proofs which are not typically mathematical may occur in mathematics too, while some mathematical reasonings may occur in non-mathematical contexts. Many combinatorial real-life situations require a mathematical approach, while games like chess require deductive thinking (although chess thinking seems to be much more than deduction). In order to identify the nature of a mathematical proof we should first delimit the idea of a ‘mathematical statement’, i.e. a statement that requires a mathematical proof. Most statements in everyday life are not of this type. Even most statements of the type ‘if . . . , then . . . ’ are not mathematical statements. At what moment does mathematics enter the scene? The answer is related to the conceptual status of the involved terms and predicates. Usually, problems raised by non-mathematicians are not yet mathematical problems, they may be farther or nearer to this status. The problem raised to Kepler, about the densest packing, in a container, of some apples of similar dimensions, was very near to a mathematical one and it was easy to find its mathematical version. The task was more difficult for the 4CP, where things like ‘map’, ‘colors’, ‘neighbor’, and ‘country’ required some delicate analysis until their mathematical models were identified. On the other hand, a question such as ‘do you love me?’ still remains far from a mathematical modelling process.

³ We have discussed in detail this issue in a previous article [11].

4 Where Does the Job of Mathematicians Begin?

Is the transition from statements in general to mathematical statements the job of mathematicians? Mathematicians are divided in answering this question. Hugo Steinhaus's answer was definitely yes, Paul Erdős's answer was clearly negative. The former liked to see in any piece of reality a potential mathematical problem, the latter liked to deal with problems already formulated in a clear mathematical language. Many intermediate situations are possible, and they give rise to a whole typology of mathematicians. Goethe's remark about mathematicians' habit of translating into their own language what you tell them and making in this way your question completely hermetic refers just to this transition, sometimes of high difficulty.

If in mathematical research both above attitudes are interesting, useful and equally important, in the field of mathematical education of the general public the yes attitude seems more important than the negative one and deserves priority. The social failure of mathematics to be recognized as a cultural enterprise is due, to a large extent, to the insufficient attention paid to its links to other fields of knowledge and creativity. This means that, in general mathematical education, besides the scenario with definitions-axioms-lemmas-theorems-proofs-corollaries-examples-applications we should consider, with at least the same attention, the scenario stressing problems, concepts, examples, ideas, motivations, the historical and cultural context, including links to other fields and ways of thinking. Are these two scenarios incompatible? Not at all. It happens that the second scenario was systematically neglected; but the historical reasons for this mistake will not be discussed here (see more in [29, 30]).

Going back to proof, perhaps the most important task of mathematical education is to explain why, in many circumstances, informal statements of problems and informal proofs are not sufficient; then, how informal statements can be translated into mathematical ones. This task is genuinely related to the explanation of what is the mathematical way of thinking, in all its variants: combinatorial, deductive, inductive, analogical, metaphorical, recursive, algorithmic, probabilistic, infinite, topological, binary, triadic, etc., and, above all, the step-by-step procedure leading to the need to use some means transcending the natural language (artificial symbols of various types and their combinations).

5 Proofs: From Pride to Arrogance

With Euclid's *Elements*, for a long time taken to be a model of rigor, mathematicians became proud of their science, claimed to be the only one giving the feeling of certainty, of complete confidence in its statements and ways of arguing. Despite some mishaps occurring in the 19th century and in the first half of the 20th century, mathematicians continued to trust in axiomatic-deductive rigor, with the improvements brought by Hilbert's ideas on axiomatics and formalization. With Bourbaki's approach, towards the middle of the 20th century, some mathematicians changed pride into arrogance, imposing a ritual excluding any

concession to non-formal arguments. ‘Mathematics’ means ‘proof’ and ‘proof’ means ‘formal proof’, is the new slogan.

Depuis les Grecs, qui dit Mathématique, dit démonstration

is Bourbaki’s slogan, while Mac Lane’s [25] austere doctrine reads

If a result has not yet been given valid proof, it isn’t yet mathematics: we should strive to make it such.

Here, the proof is conceived according to the standards established by Hilbert, for whom a proof is a demonstrative text starting from axioms and where each step is obtained from the preceding ones, by using some pre-established explicit inference rules:

The rules should be so clear, that if somebody gives you what they claim is a proof, there is a mechanical procedure that will check whether the proof is correct or not, whether it obeys the rules or not.

And according to Jaffe and Quinn [20]

Modern mathematics is nearly characterized by the use of rigorous proofs. This practice, the result of literally thousands of years of refinement, has brought to mathematics a clarity and reliability unmatched by any other science.

This is a *linear-growth model* of mathematics (see Stöltzner [36]), a process in two stages. First, informal ideas are guessed and developed, conjectures are made, and outlines of justifications are suggested. Secondly, conjectures and speculations are tested and corrected; they are made reliable by *proving* them. The main goal of proof is to provide reliability to mathematical claims. The act of finding a proof often yields, as a by-product, new insights and possibly unexpected new data. So, by making sure that every step is correct, one can tell once and for all whether a theorem has been proved. Simple! A moment of reflection shows that the case may not be so simple. For example, what if the “agent” (human or computer) checking a proof for correctness makes a mistake (as pointed out by Lakatos [24], agents are fallible)? Obviously, another agent has to check that the agent doing the checking did not make any mistake. Some other agent will need to check that agent, and so on. Eventually either the process continues unendingly (an unrealistic scenario?), or one runs out of agents who could check the proof and, in principle, they could all have made a mistake! Finally, the linear-growth model is built on an asymmetry of proof and conjecture: Posing the latter does not necessarily involve proof.

The Hilbert-Bourbaki model has its own critics, some from outside mathematics such as Lakatos [24]

... those who, because of the usual deductive presentation of mathematics, come to believe that the path of discovery is from axiom and/or definitions to proofs and theorems, may completely forget about the possibility and importance of naive guessing

some from eminent mathematicians as Atiyah [2]:

[20] *present a sanitized view of mathematics which condemns the subject to an arthritic old age. They see an inexorable increase in standards and are embarrassed by earlier periods of sloppy reasoning. But if mathematics is to rejuvenate itself and break new ground it will have to allow for the exploration of new ideas and techniques which, in their creative phase, are likely to be dubious as in some of the great eras of the past. Perhaps we now have high standards of proof to aim at but, in the early stages of new developments, we must be prepared to act in more buccaneering style.*

Atiyah's point meets Lakatos's [24] views

... informal, quasi-empirical mathematics does not grow through a monotonous increase of the number of indubitably established theorems, but through the incessant improvement of guesses by speculation and criticism, by the logic of proof and refutation

and is consistent with the idea that the linear-growth model tacitly requires a 'quasi-empirical' ontology, as noted by Hirsch in his contribution to the debate reported in [2]:

For if we don't assume that mathematical speculations are about 'reality' then the analogy with physics is greatly weakened—and there is no reason to suggest that a speculative mathematical argument is a theory of anything, any more than a poem or novel is 'theoretical'.

6 Proofs: From Arrogance to Prudence

It is well-known that the doubt appeared in respect to the Hilbert-Bourbaki rigor was caused by Gödel's incompleteness theorem,⁴ see, for instance, Kline's *Mathematics, the Loss of Certainty* [21]. It is not by chance that a similar title was used later by Ilya Prigogine in respect to the development of physics. So, arrogance was more and more replaced by prudence. All rigid attitudes, based on binary predicates, no longer correspond to the new reality, and they should be considered 'cum grano salis'. The decisive step in this respect was accomplished by the spread of empirical-experimental factors in the development of proofs.

⁴ The result has generated a variety of reactions, ranging from pessimism (the final, definite failure of any attempt to formalise all of mathematics) to optimism (a guarantee that mathematics will go on forever) or simple dismissal (as irrelevant for the practice of mathematics). See more in Barrow [4], Chaitin [12] and Rozenberg and Salomaa [34]. The main pragmatical conclusion seems to be that 'mathematical knowledge', whatever this may mean, cannot solely be derived only from some fixed rules. Then, who validates the 'mathematical knowledge'? Wittgenstein's answer was that the *acceptability ultimately comes from the collective opinion of the social group of people practising mathematics.*

7 Assisted Proofs Vs. Long Proofs, or from Prudence to Humility

The first major step was realized in 1976, with the discovery, using a massive computer computation, of a proof of the 4CP. This event should be related to another one: the increasing length of some mathematical proofs. Obviously, the length $l(p(s))$ of the proof $p(s)$ of the statement s should be appreciated in respect to the length $l(s)$ of s . There is a proposal to require the existence of a strictly positive constant k such that, for any reasonable theorem, the ratio $l(p(s))/l(s)$ is situated between $1/k$ and k . But the existence of such a k may remain an eternal challenge.

In the past, theorems with too long a statement were very rare. Early examples of this type can be found in Apollonius's *Conica* written some time after 200 BC. More recent examples include some theorems by Arnaud Denjoy, proved in the first decades of the 20th century, and Jordan's theorem (1870) concerning the way a simple closed curve c separates the plane in two domains whose common frontier is c . A strong trend towards long proofs appears in the second half of the 20th century. We exclude here the artificial situation when theorems with long statements and long proofs can be decomposed into several theorems, with normal lengths. We refer to statements having a clear meaning, whose unity and coherence are lost if they are not maintained in their initial form. The 4CP is just of this type. Kepler's conjecture is of the same type and so are Fermat's theorem, Poincaré's conjecture and Riemann's hypothesis. What about the theorem giving the classification of finite simple groups? In contrast with the preceding examples, in this case the statement of the theorem is very long. It may be interesting to observe that some theorems which are in complete agreement with our intuition, like Jordan's and Kepler's, require long proofs, while some other theorems, in conflict with our intuition, such as the theorem asserting the existence of three domains in the plane having the same frontier, have a short proof. Ultimately, everything depends on the way the mathematical text is segmented in various pieces.

The proof of the theorem giving the typology of the finite simple groups required a total of about fifteen thousand pages, spread in five-hundred separate articles belonging to about three-hundred different authors (see Conder [15]). But Serre [31] is still waiting for experts to check the claim by Aschbacher and Smith to have succeeded filling in the gap in the proof of the classification theorem, a gap already discovered in 1980 by Daniel Gorenstein. The gap concerned that part which deals with 'quasi-thin' groups. Despite this persisting doubt, most parts of the global proof were already published in various prestigious journals. The ambition of rigor was transgressed by the realities of mathematical life. Moreover, while each author had personal control of his own contribution (excepting the mentioned gap), the general belief was that the only person having a global, holistic representation and understanding of this theorem was Daniel Gorenstein, who unfortunately died in 1992. So, the classification theorem is still looking for its validity and understanding.

The story of the classification theorem points out the dramatic fate of some mathematical truths, whose recognition may depend on sociological factors which are no longer under the control of the mathematical community. This situation is not isolated. Think of Fermat's theorem, whose proof (by Wiles) was checked by a small number of specialists in the field, but the fact that here we had several 'Gorensteins', not only one, does not essentially change the situation.

How do exceedingly long proofs compare with assisted proofs? In 1996 Robertson, Sanders, Seymour and Thomas [32] offered a simpler proof of the 4CP. They conclude with the following interesting comment (p. 24):

*We should mention that both our programs use only integer arithmetic, and so we need not be concerned with round-off errors and similar dangers of floating point arithmetic. However, an argument can be made that our "proof" is not a proof in the traditional sense, because it contains steps that can never be verified by humans. In particular, we have not proved the correctness of the compiler we compiled our programs on, nor have we proved the infallibility of the hardware we ran our programs on. These have to be taken on faith, and are conceivably a source of error. However, from a practical point of view, the chance of a computer error that appears consistently in exactly the same way on all runs of our programs on all the compilers under all the operating systems that our programs run on is infinitesimally small compared to the chance of a human error during the same amount of case-checking. Apart from this hypothetical possibility of a computer consistently giving an incorrect answer, the rest of our proof can be verified in the same way as traditional mathematical proofs. We concede, however, that verifying a computer program is much more difficult than checking a mathematical proof of the same length.*⁵

Knuth [22] p. 18 confirms the opinion expressed in the last lines of the previous paragraph:

... program-writing is substantially more demanding than book-writing. Why is this so? I think the main reason is that a larger attention span is needed when working on a large computer program than when doing other intellectual tasks. ... Another reason is ... that programming demands a significantly higher standard of accuracy. Things don't simply have to make sense to another human being, they must make sense to a computer.

And indeed, Knuth compared his \TeX compiler (a document of about 500 pages) with Feit and Thompson's [17] theorem that all simple groups of odd order are cyclic. He lucidly argues that the program might not incorporate as much creativity and "daring" as the proof of the theorem, but they come even when compared on depth of detail, length and paradigms involved. What distinguishes the program from the proof is the "verification": convincing a couple of (human) experts that the proof *works in principle* seems to be easier than making sure that

⁵ Our emphasis.

the program *really works*. A demonstration that *there exists a way to compile T_EX* is not enough! Hence Knuth's warning: "Beware of bugs in the above code: I have only proved it correct, not tried it."

It is just the moment to ask, together with R. Graham: "If no human being can ever hope to check a proof, is it really a proof?" Continuing this question, we may ask: What about the fate of a mathematical theorem whose understanding is in the hands of only a few persons? Let us observe that in both cases discussed above (4CP and the classification theorem) it is not only the global, holistic understanding under question, but also its local validity.

Another example of humility some eminent mathematicians are forced to adopt with respect to yesterday's high exigency of rigor was given recently by one of the most prestigious mathematical journals, situated for a long time at the top of mathematical creativity: *Annals of Mathematics*. We learn from Karl Sigmund [35] that the proof proposed by Thomas Hales in August 1998 and the corresponding joint paper by Hales and Ferguson confirming Kepler's conjecture about the densest possible packing of unit spheres into a container, was accepted for publication in the *Annals of Mathematics*,

but with an introductory remark by the editors, a disclaimer as it were, stating that they had been unable to verify the correctness of the 250-page manuscript with absolute certainty.

The proof is so long and based to such an extent on massive computations, that the platoon of mathematicians charged with the task of checking it ran out of steam. Robert MacPherson, the *Annals'* editor in charge of the project, stated that "the referees put a level of energy into this that is, in my experience, unprecedented. But they ended up being only 99 percent certain that the proof was correct". However, not only the referees, the author himself, Thomas Hales, 'was exhausted', as Sigmund observes. He was advised to re-write the manuscript: he didn't, but instead he started another project, 'Formal Proof of Kepler' (FPK), a project which puts theorem-verification on equal footing with Knuth's program-verification. Programming a machine to check human reasoning gives a new type of insight which has its own kind of beauty. Here is the bitter-ironical comment by Sigmund:

After computer-based theorem-proving, this is the next great leap forward: computer-based proof checking. Pushed to the limit, this would seem to entail a self-referential loop. Maybe the purists who insist that a proof is a proof if they can understand it are right after all. On the other hand, computer-based refereeing is such a promising concept, for reviewers, editors, and authors alike, that it seems unthinkable that the community will not succumb to the temptation.

So, what is the perspective? It appears that FPK will require 20 man-years to check every single step of Hales' proof. "If all goes well, we then can be 100 percent certain", concludes Sigmund ([35], p. 67).

Let us recall that Perelman's recent proof of Poincaré's conjecture⁶ is still being checked at MIT (Cambridge) and IHES (Paris) and who knows how long this process will be? We enter a period in which mathematical assessment will increase in importance and will use, in its turn, computational means. The job of an increasing number of mathematicians will be to check the work of other mathematicians. We have to learn to reward this very difficult work, to pay it at its correct value.

One could think that the new trend fits the linear-growth model: all experiments, computations and simulations, no matter how clever and powerful, belong to and are to stay at the first stage of mathematical research where informality and guessing are dominant. *This is not the case.* Of course, some automated heuristics will belong only to the first stage. The shift is produced when a large part of the results produced by computing experiments are transferred to the second stage; they no longer only develop the intuition, they no longer only build hypotheses, but *they assist the very process of proof, from discovery to checking, they create a new type of environment in which mathematicians can undertake mathematical research.*

For some a proof including computer programs is like a telephone directory, while a human proof may compete with a beautiful novel. This analogy refers to the exclusive syntactic nature of a computer-based proof (where we learn that the respective proof is valid, but we may not (don't) understand why), contrasting with the attention paid to the semantic aspect, to the understanding process, in the traditional proofs, exclusively made by humans.⁷ The criticism implied by this analogy, which is very strong in René Thom's writings, is not always motivated. In fact, an 'elegant' program⁸ may help the understanding process of mathematical facts in a completely new way. We confine ourselves to a few examples only:

... if one can program a computer to perform some part of mathematics, then in a very effective sense one does understand that part of mathematics (G. Tee [38])

If I can give an abstract proof of something, I'm reasonably happy. But if I can get a concrete, computational proof and actually produce numbers I'm much happier. I'm rather an addict of doing things on computer, because that gives you an explicit criterion of what's going on. I have a visual way of thinking, and I'm happy if I can see a picture of what I'm working with (J. Milnor, [7])

... computer-based proofs are often more convincing than many standard proofs based on diagrams which are claimed to commute, arrows which

⁶ Mathematicians familiar with Perelman's work expect that it will be difficult to locate any substantial mistakes, cf. Robinson [33].

⁷ The conjugate pair rigor-meaning deserves to be reconsidered, cf. Marcus [26].

⁸ Knuth's concept of treating a program as a piece of literature, addressed to human beings rather than to a computer; see [23].

are supposed to be the same, and arguments which are left to the reader (J.-P. Serre [31])

... the computer changes epistemology, it changes the meaning of "to understand." To me, you understand something only if you can program it. (G. Chaitin [14])

It is the right moment to *reject the idea that computer-based proofs are necessarily ugly and opaque not only to being checked for their correctness, but also to being understood in their essence.*

Finally, do axiomatic-deductive proofs remain an *a posteriori* work, a luxury we can marginalize? When asked whether "when you are doing mathematics, can you know that something is true even before you have the proof?", Serre ([31], p. 212) answers: "Of course, this is very common". But he adds: "But one should distinguish between the genuine goal [...] which one feels is surely true, and the auxiliary statements (lemmas, etc.), which may well be intractable (as happened to Wiles in his first attempt) or even downright false [...]."

8 A Possible Readership Crisis and the Globalization of the Proving Process

Another aspect of very long (human or computer-assisted) proofs is the risk of finding no competent reader for them, no professional mathematician ready to spend a long period to check them. This happened with the famous Bieberbach conjecture. In 1916, L. Bieberbach conjectured a necessary condition on an analytic function to map the unit disk injectively to itself. The statement concerns the (normalised) Taylor coefficients a_n of such a function ($a_0 = 0, a_1 = 1$): it then states that $|a_n|$ is at most n , for any positive integer n . Various mathematicians succeeded in proving the required inequality for particular values of n , but not for every n . In March 1984, Louis de Branges (from Purdue University, Lafayette) claimed a proof, but nobody trusted him, because previously he made wrong claims for other open problems. Moreover, nobody in USA agreed to read his 400-pages manuscript to check his proof, representing seven years of hard work. The readership crisis ended when Louis de Branges proposed to the Russian mathematician I. M. Milin that he check the proof; Milin was the author of a conjecture implying Bieberbach's conjecture. De Branges travelled to Leningrad, where after a period of three months of confrontation with a team formed by Milin and two other Russian mathematicians, E. G. Emelianov and G. V. Kuzmina, they all reached the conclusion that various mistakes existing in the proof were all benign. Stimulated by this fact, two German mathematicians, C. F. Gerald and C. Pommerenke (Technical University, Berlin) succeeded in simplifying De Branges's proof.

This example is very significant for the globalization of mathematical research, a result of the globalization of communication and of international co-operation. It is no exaggeration to say that *mathematical proof has now a global dimension.*

9 Experimental Mathematics or the Hope for It

The emergence of powerful mathematical computing environments such as Mathematica, MathLab, or Maple, the increasing availability of powerful (multi-processor) computers, and the omnipresence of the Internet allowing mathematicians to proceed heuristically and ‘quasi-inductively’, have created a blend of logical and empirical–experimental arguments which is called “quasi–empirical mathematics” (by Tymoczko [39], Chaitin [13]) or “experimental mathematics” (Borwein, Bailey [8], Borwein, Bailey, Girgensohn [9]). Mathematicians increasingly use symbolic and numerical computation, visualisation tools, simulation and data–mining. New types of proofs motivated by the experimental “ideology” have appeared. For example, the *interactive proof* (see Goldwasser, Micali, Rackoff [18], Blum [5]) or the *holographic proof* (see Babai [3]). And, of course, these new developments have put the classical idea of axiomatic–deductive proof under siege (see [11] for a detailed discussion).

Two programmatic ‘institutions’ are symptomatic for the new trend: the *Centre for Experimental and Constructive Mathematics (CECM)*,⁹ and the journal *Experimental Mathematics*.¹⁰ Here are their working ‘philosophies’:

At CECM we are interested in developing methods for exploiting mathematical computation as a tool in the development of mathematical intuition, in hypothesis building, in the generation of symbolically assisted proofs, and in the construction of a flexible computer environment in which researchers and research students can undertake such research. That is, in doing experimental mathematics. [6]

Experimental Mathematics publishes formal results inspired by experimentation, conjectures suggested by experiments, surveys of areas of mathematics from the experimental point of view, descriptions of algorithms and software for mathematical exploration, and general articles of interest to the community.

For centuries mathematicians have used experiments, some leading to important discoveries: the Gibbs phenomenon in Fourier analysis, the deterministic chaos phenomenon, fractals. Wolfram’s extensive computer experiments in theoretical physics paved the way for his discovery of simple programs having extremely complicated behavior [40]. Experimental mathematics—as systematic mathematical experimentation ranging from hypotheses building to assisted proofs and automated proof–checking—will play an increasingly important role and will become part of the mainstream of mathematics. There are many reasons for this trend: they range from logical (the absolute *truth* simply doesn’t exist), sociological (correctness is not absolute as mathematics advances by making mistakes and correcting and re–correcting them), economic (powerful computers will be accessible to more and more people), and psychological (results and success

⁹ www.cecm.sfu.ca.

¹⁰ www.expmath.org.

inspire emulation). The computer is *the* essential, but not the *only* tool. New theoretical concepts will emerge, for example, the systematic search for new axioms. Assisted-proofs are not only useful and correct, but they have their own *beauty* and *elegance*, impossible to find in classical proofs. The experimental trend is not antithetical to the axiomatic-deductive approach, it complements it. Nor is the axiomatic-deductive proof a *posteriori* work, a luxury we can marginalize. There is hope for integration!

Acknowledgment

We are grateful to Greg Chaitin and Garry Tee for useful comments and references.

References

1. K. Appel, W. Haken. *Every Planar Graph is Four Colorable*, Contemporary Mathematics 98, AMS, Providence, 1989.
2. M. Atiyah et al. Responses to ‘Theoretical mathematics: Toward a cultural synthesis of mathematics and theoretical physics’, *Bulletin of AMS* 30 (1994), 178–211.
3. L. Babai. Probably true theorems, cry wolf? *Notices of AMS* 41 (5) (1994), 453–454.
4. J. Barrow. *Impossibility—The Limits of Science and the Science of Limits*, Oxford University Press, Oxford, 1998.
5. M. Blum. How to prove a theorem so no one else can claim it, *Proceedings of the International Congress of Mathematicians*, Berkeley, California, USA, 1986, 1444–1451.
6. J. M. Borwein. Experimental Mathematics and Integer Relations at www.ercim.org/publication/Ercim_News/enw50/borwein.html.
7. J. M. Borwein. www.cecm.sfu.ca/personal/jborwein/CRM.html.
8. J. M. Borwein, D. Bailey. *Mathematics by Experiment: Plausible Reasoning in the 21st Century*, A.ËK.Ë Peters, Natick, MA, 2003.
9. J. M. Borwein, D. Bailey, R. Girgensohn. *Experimentation in Mathematics: Computational Paths to Discovery*, A.ËK.Ë Peters, Natick, MA, 2004.
10. A. S. Calude. The journey of the four colour theorem through time, *The NZ Math. Magazine* 38, 3 (2001), 27–35.
11. C. S. Calude, E. Calude, S. Marcus. Passages of Proof, Los Alamos preprint archive, [arXiv:math.HO/0305213](https://arxiv.org/abs/math/0305213), 16 May 2003.
12. G. J. Chaitin. *The Unknowable*, Springer Verlag, Singapore, 1999.
13. G. J. Chaitin. *Exploring Randomness*, Springer Verlag, London, 2001.
14. G. J. Chaitin. *Meta Math!*, E-book at www.cs.auckland.ac.nz/CDMTCS/chaitin/omega.html.
15. M. Conder. Pure mathematics: An art? or an experimental science? *NZ Science Review* 51, 3 (1994), 99–102.
16. E. Ehrlich, S. B. Flexner, G. Carruth, J. M. Hawkins. *Oxford American Dictionary*, Avon Publishers of Bard, Camelot, Discus and Flare Books, New York, 1982.
17. W. Feit, J. G. Thomson. Solvability of groups of odd order, *Pacific J. Math.* 13 (1963), 775–1029.

18. S. Goldwasser, S. Micali, C. Rackoff. The knowledge complexity of interactive proof-systems, *SIAM J. Comput.*, 18(1) (1989), 186–208.
19. R. Hersh. *What Is Mathematics, Really?*, Vintage, London, 1997.
20. A. Jaffe and F. Quinn. Theoretical mathematics: Toward a cultural synthesis of mathematics and theoretical physics, *Bulletin of AMS* 29 (1993), 178–211.
21. M. Kline. *Mathematics: The Loss of Certainty*, Oxford University Press, Oxford, 1982.
22. D. E. Knuth. Theory and practice, *EATCS Bull.* 27 (1985), 14–21.
23. D. E. Knuth. *Literate Programming*, CSLI Lecture Notes, no. 27, Stanford, California, 1992.
24. I. Lakatos. *Proofs and Refutations. The Logic of Mathematical Discovery*, John Worrall and Elie Zahar (eds.), Cambridge University Press, Cambridge, 1966.
25. S. Mac Lane. Despite physicists, proof is essential in mathematics, *Synthese* 111 (1997), 147–154.
26. S. Marcus. No system can be improved in all respects, in G. Altmann, W. Koch (eds.) *Systems; New Paradigms for the Human Sciences*, Walter de Gruyter, Berlin, 1998, 143–164.
27. S. Marcus. *Ways of Thinking*, Scientific and Encyclopedic Publ. House, Bucharest, 1987. (in Romanian)
28. C. Perelman, L. Olbrechts-Tyteca. *Traité de l'Argumentation. La Nouvelle Rhétorique*, Éditions de l'Université de Bruxelles, Bruxelles, 1988.
29. G. Pólya. *How to Solve It*, Princeton University Press, Princeton, 1957. (2nd edition)
30. G. Pólya. *Mathematics and Plausible Reasoning*, Volume 1: *Induction and Analogy in Mathematics*, Volume 2: *Patterns of Plausible Inference*, Princeton University Press, Princeton, 1990. (reprint edition)
31. M. Raussen, C. Skau. Interview with Jean-Pierre Serre, *Notices of AMS*, 51, 2 (2004), 210–214.
32. N. Robertson, D. Sanders, P. Seymour, R. Thomas. A new proof of the four-colour theorem, *Electronic Research Announcements of AMS* 2,1 (1996), 17–25.
33. S. Robinson. Russian reports he has solved a celebrated math problem, *The New York Times*, April 15 (2003), p.ÉD3.
34. G. Rozenberg, A. Salomaa. *Cornerstones of Undecidability*, Prentice-Hall, New York, 1994.
35. K. Sigmund. Review of George G. Szpiro. “Kepler’s Conjecture”, Wiley, 2003, *Mathematical Intelligencer*, 26, 1 (2004), 66–67.
36. M. Stöltzner. What Lakatos could teach the mathematical physicist, in G. Kampis, L. Kvasz, M. Stöltzner (eds.). *Appraising Lakatos. Mathematics, Methodology and the Man*, Kluwer, Dordrecht, 2002, 157–188.
37. E. R. Swart. The philosophical implications of the four-colour problem, *American Math. Monthly* 87, 9 (1980), 697–702.
38. G. J. Tee. Computers and mathematics, *The NZ Math. Magazine* 24, 3 (1987), 3–9.
39. T. Tymoczko. The four-colour problem and its philosophical significance, *J. Philosophy* 2,2 (1979), 57–83.
40. S. Wolfram. *A New Kind of Science*, Wolfram Media, 2002.
41. *Experimental Mathematics: Statement of Philosophy*, www.expmath.org/expmath/philosophy.html.