3

On the Roles of Proof in Mathematics

Joseph Auslander
Department of Mathematics
University of Maryland

From the Editors

This third perspective on proof comes from a mathematician with a more traditional perspective than Borwein's. The author brings his considerable experience both in developing his own proofs and in reviewing others' to questions about the roles of proof. His discussion on the roles of proof contains some interesting new ideas, such as proof as exploration and proof as justification of definitions—ideas that are relevant to us as we think about how we teach mathematics. At the end he offers some extended illustrations of his main points, from his experience working in topological dynamics and ergodic theory.

Joseph Auslander is a Professor Emeritus of Mathematics at the University of Maryland. He has published extensively in topological dynamics and ergodic theory. He is the author of Minimal Flows and Their Extensions *(1988) and co-editor, with Walter H. Gottschalk, of* Topological Dynamics, *an international symposium (1968). He has published two reviews of books in the philosophy of mathematics:* What is Mathematics, Really? *by Reuben Hersh,* Where Mathematics Comes From: How the Embodied Mind Brings Mathematics into Being *by George Lakoff and Rafael E. Núñez. Those reviews appeared in* SIAM Review *(2000) and* American Scientist *(2001), respectively. With Bonnie Gold, he organized a panel for the winter 2001 joint mathematics meetings in New Orleans on "The Philosophy of Mathematics: That Which is of Interest to Mathematicians," which led to the founding of POMSIGMAA. He was the first Secretary of POMSIGMAA, and gave a talk, "When is a Proof a Proof?" at the POMSIGMAA contributed paper session in January 2004.*

In this article, I will make, and try to justify, the following points.

Deductive proof is almost the defining feature of mathematics. Mathematics without proof would not be mathematics. This is so although mathematics consists of more than proof, and proof occurs in other disciplines.

Proof is necessary for validation of a mathematical result. But there are other, equally compelling reasons for proof.

Standards of proof vary over time, and even among different mathematicians at a given time.

The question of "when is a proof a proof?" is a complex one. This has always been an issue, but it is particularly so now in the light of computer assisted proofs and very long proofs.

1 Proof as a Defining Feature of Mathematics

I am writing as a working mathematician, not as a philosopher. My approach to proof is consistent with the viewpoint, cogently put forth by Reuben Hersh [1997] and Paul Ernest [1998], that mathematics is socially constructed. That is, it has been constructed by humans, and is part of human culture. Therefore I will focus on what mathematicians actually do. This is what Hersh calls "practical proof—the argument that convinces the qualified skeptical expert" rather than formal proof.

Thomas Hales clarifies this distinction well [Hales www.]:

> "Traditional mathematical proofs are written in a way to make them easily understood by mathematicians. Routine logical steps are omitted. An enormous amount of context is assumed on the part of the reader. Proofs, especially in topology and geometry, rely on intuitive arguments in situations where a trained mathematician would be capable of translating those intuitive arguments into a more rigorous argument." This is distinguished from formal proof where "all the intermediate logical steps are supplied" and "no appeal is made to intuition."

I will not try to give a precise definition of mathematics; the definitions I've seen are either too restrictive or too inclusive, but certainly the use of deductive proof is an essential feature. Mathematics is not just about "results." (One might refer to the belief in the primacy of results to the exclusion of anything else as the "Vince Lombardi" approach, after the football coach who said that "winning is the only thing.")

Mathematics is a process, which includes definitions, conjectures, examples, numerical evidence, statements of theorems, modelling, algorithms, and proofs, as well as heuristic arguments which fall short of proof. These are all woven together. In particular the proof is inextricably bound up with the result; indeed one can't really separate them. This is part of the aesthetics of mathematics, but it also has "practical" consequences. Proofs often contain "subresults," as well as implicit or explicit lemmas, which are of interest in themselves. These would be lost if one just catalogued "results." Moreover, often a proof yields more than is explicitly stated, and it may point the way to new theorems. This is illustrated by Hillel Furstenberg's proof of the Szemeredi theorem, which will be discussed below.

As John Franks [1989] eloquently puts it "a proof is not some kind of super spell checker that merely validates mathematical facts... Proofs (are) the central content of mathematical knowledge... Who would be satisfied if God were to announce that the Riemann hypothesis is true, but deny us the proof?" (Regarding the last point, we might ask if we would be satisfied if a computer "announced" that a theorem had been proved, but we couldn't see the proof. See the section on "Proof as Certification" for more about this.)

Another gloss on this topic was stated by the biologist Richard Lewontin [2005] writing in the New York Review of Books: "Science, indeed scholarship in general, is a domain in which the integrity of the process is more important than . . . any particular result. This is . . . a question of the very survival of the process of investigation." Lewontin in this passage was concerned with issues of honesty and fraud in science, but the point holds in a more general context.

As was mentioned above, mathematics is not only about proof. Moreover, the notion of proof also occurs in other areas (in other sciences of course—physical, biological, and social—and also such disciplines as law and history) but it has a somewhat different meaning, and different methods are used to attain it. These are characterized by a mixture of deductive reasoning and empirical evidence. Debates about the relation of these are at the heart of the philosophy of science.

I should say at the outset that I am definitely not asserting that proofs in mathematics are in some sense more "valid" than those in other disciplines. Rather, there are different methods of arriving at conclusions, and that deductive proof is central to mathematics to a much higher degree than in other areas. This is in spite of some challenges to this central role and even some predictions of the "death of proof."

We might accept as a provisional definition of proof a (valid) sequence of deductions, starting with the hypothesis, and arriving at the conclusion. Somewhat more formally [Kitcher 1984, p. 38] "We can now define a proof as a sequence of statements such that every member of the sequence is either a basic *a priori* statement or a statement which follows from previous members of the sequence in accordance with some apriority-preserving rule of inference."

This is somewhat at variance with our earlier emphasis on "practical proof," and in fact this tension is one of the things that makes the issue interesting. Nevertheless, mathematicians do feel that their proofs essentially accomplish what Kitcher describes (as the quotation from Hales in Section 1 points out). In fact, it's fair to say that this is a necessary and sufficient condition for a proof. That is, if this is achieved, we have a proof, and if it isn't there is no proof.

In a sense, that's all there is to it. As Gian-Carlo Rota [1996] puts it, "Mathematical proof does not admit degrees. A sequence of steps in an argument is either a proof, or it is meaningless. . . . The mathematical notion of proof is strikingly at variance with notions of proof in . . . law, everyday conversation, and physics."

However, I will argue that the situation is more complex than Rota makes it out to be. While any two mathematicians will agree in the abstract what a proof is, it's when one gets down to cases that problems may arise. Many of these can be reduced to "how do we know that a theorem has in fact been proved?" For example, what about "Proof: Obvious," or "Proof: This follows from the previous lemma?" At the other extreme, how do we evaluate a 15000 page proof, which may itself rely on papers the author hasn't read? Or a proof dependent on an unpublished or unobtainable paper? Or, a hot topic these days, a proof making use of computer calculations?

The issue of proofs in elementary and secondary school mathematics has been much discussed. Many (this writer included) lament the lack of emphasis on proofs in today's high school geometry classes, in contrast to what occurred in previous generations. The following quotation of Ken Ross [1998] addresses this point.

> "While science verifies through observation, mathematics verifies through logical reasoning. Thus the essence of mathematics lies in proofs. . . . It should be emphasized that results in mathematics follow from hypotheses. . . . Moreover, beginning in the 8th

grade, students should distinguish between inductive and deductive reasoning, be able to identify the hypothesis and conclusion in a deduction, test an assertion with examples, realize that one counterexample is enough to show that an assertion is false, and recognize whether something is being proved or merely given a plausibility argument."

2 The Roles of Proof

Mathematicians have a range of views on the role of proof in mathematics. Several of these views are illustrated by the following quotations, in which I have italicized words that emphasize the role of proof being mentioned.

Hyman Bass [2003]: "The characteristic that distinguishes mathematics from all other sciences is the nature of mathematical knowledge and its *certification* by means of mathematical proof... it is the only science that thus pretends to claims of absolute certainty."

Gian-Carlo Rota [1993, p. 93]: "Mathematicians cannot afford to behave like physicists, who take experimental verification as *confirmation* of the truth."

In fact, the physicist Steven Weinberg [2001] makes essentially the same point as Rota: "You give up worrying about *certainty* when you make that turn in your career that makes you a physicist rather than a mathematician."

David Gale [1990]: "The main goal of science is to observe and then to explain phenomena. In mathematics the *explanation* is the proof... the theorem-proof methodology... (is) the only *methodology* we have."

Philip Davis and Reuben Hersh [1981, p. 151]: "Proof serves many purposes simultaneously... (It is) subject to a constant process of criticism and revalidation. Errors, ambiguities, and misunderstandings are cleared up by constant exposure. Proof is respectability. Proof is the seal of authority... (It) increases understanding by revealing the heart of the matter. Proof suggests new mathematics. Proof is mathematical power, the electric voltage of the subject which vitalizes the static assertions of the theorems."

Saunders Mac Lane [Responses 1994, p.190]: "Intuition is glorious, but the heaven of mathematics requires much more.... Mathematics rests on proof—and proof is eternal."

There is no doubt that the overwhelming majority of mathematicians is committed to proofs in the traditional sense, and endorses the sentiments, if not the exact wording of the above quotations. Later, I'll express reservations about some of the assertions.

I would like to single out several (not unrelated) roles of proof, including certification (or validation), explanation, and exploration.

2.1 Proof as Certification

We accept that a purported result is correct when we hear that it has been proved by a mathematician we trust and "validated" by experts in the author's mathematical specialty. This is the case even if we haven't read the proof, or more frequently when we don't have the background to follow the proof. As an extreme, perhaps hackneyed, example, mathematicians accept Wiles' proof of Fermat's last theorem because number theorists have "certified" it to be correct. While certification is the most "primitive" or "elementary" aspect of proof, it is worthwhile looking at this role more closely. It is an indication that we are part of a community whose members trust one another. In fact, mathematics could not be a coherent discipline, as opposed to a random collection of techniques and results, without the process of certification.

Usually, certification of a result is a consequence of its appearance as a paper in a refereed journal. In fact, we might agree that this is a necessary condition for certification. In this case it is generally accepted that the "burden of proof" (the pun is inevitable) has shifted, and the result is presumed correct, unless there is a compelling reason to believe otherwise. It should be emphasized that it's necessary that one is convinced that a competent mathematician has worked out the proof, rather than it being "announced" by "God" (as in the quotation from John Franks earlier).

However, this process is far from perfect, and should be regarded as provisional. For one thing, it is well known that standards of refereeing vary widely. Some papers—for example, the proof of Fermat's last theorem, and Hales' proof of the Kepler conjecture discussed below—concern famous problems, and thus have received intense scrutiny. Other papers receive more routine treatment. Ralph Boas, who was for many years the editor of *Mathematical Reviews*, is said to have remarked that of the new results in papers reviewed most are true but the corresponding proofs are perhaps half the time wrong.

An interesting example was the published assertion by Waraskiewicz [1937], that a homogeneous plane continuum is necessarily a simple closed curve. This "result" was generally accepted, and in fact a more general assertion was published by Choquet [1944]. However, a counterexample was provided by Bing [1948]. (Another example will be discussed in the section "Four Examples.")

Also, referees are generally told that it is not their job to determine whether a paper is correct—this is the responsibility of the author—although the referee should be reasonably convinced. The referee is typically asked to determine whether the paper is worthwhile. Of course this begs the question somewhat. If the result is not correct, then the paper is not worthwhile. In the case of very long papers, referees usually don't try to check every line. Robert MacPherson, an editor of the *Annals of Mathematics* says "I try to understand the internal logic of the proof and do consistency checks." [Szpiro 2003, p. 208] Moreover, there are (presumably refereed) papers in respectable journals where the claimed result is false (in some cases not so noted for many years).

The issue of the refereeing process—real and ideal—in mathematics is fascinating and largely unexplored. Gossip on this topic abounds but I know of no systematic study.

Certification of a result allows us to use it in further research. In theory, one just checks the hypotheses, and if they are appropriate to the given situation, applies the result and goes on from there. This may be necessary (one can't develop all of mathematics each time one writes out a proof) but it brings along certain dangers. For reasons which aren't entirely clear, applying a result mechanically, without an understanding of the proof, can lead to errors. For example, sometimes one is fooled by notation. (This is borne out by my own experience. In fact, on one occasion I was attempting to apply something I had proved earlier without thinking it through carefully, and I made an elementary error.)

The point is that a mathematician is not absolved from understanding the proof, even when the result in question has been accepted by the mathematical community. When one uses a result in one's own research or teaching, the stakes are higher. It then becomes necessary to understand at least the basic outlines of the proof. One requires a higher degree of certainty for the use of a result than is obtained by the passive acceptance of it.

This was put well by Daniel Biss [2004]: "No honest mathematician uses a result simply because it has been published. Rather we use results we trust are true . . . the defining threshold for this notion is . . . a complex mélange of what has been published, what has been accepted as true by a larger community, and . . . what we believe ourselves to understand."

There is a recent tendency for (some) mathematicians to post their papers on preprint servers. Frequently this is preliminary to the submission of these papers to a journal (in which case it's not particularly different from the former practice of the distribution of preprints, allowing access of the results to researchers in the field prior to publication), but in some cases there is no intention of submission to a journal. Even given the imperfect process of refereeing, this somewhat undermines the certification of the results in question.

2.2 Proof as Explanation

Our second role of proof is explanation. This is what concerns most mathematicians. One should be able to follow at least the broad outlines of the argument, and be confident that one can fill in the details. As Andrew Gleason [Yandell 2001, p. 150] points out, "Proofs really aren't there to convince you that something is true . . . they're there to show you why it is true."

Ideally this is what proof is all about. Almost by definition, a proof is supposed to explain the result. Now, it must be admitted that not all proofs meet this standard. To some extent this is in the eye of the beholder. Indeed sometimes the conviction that a result is correct may arise not from the proof, but from (say) numerical evidence, illuminating examples, or visual representation. Such considerations have often led to the development of new, more understandable, proofs.

The great mathematician Paul Erdős spoke of "The Book" in which "God" maintained the "perfect" proofs of theorems. In fact there is a real book, appropriately titled *Proofs from THE BOOK* by Martin Aigner and Gunter Ziegler [1999] which presents many proofs in this spirit. Erdős collaborated on this book shortly before his death, and many of the proofs are due to him.

The first chapter consists of six different proofs of the infinity of primes, starting with the familiar proof due to Euclid. The sixth proof, due to Erdős, proves more, namely that the sum of the reciprocals of the primes diverges. (The first proof of this fact was given by Euler.) Erdős' proof is by contradiction—suppose the sum converges. If $p_1, p_2, \ldots$ is the sequence of primes written in increasing order, then there is a k such that $\sum_{i \geq k+1} \frac{1}{p_i} < \frac{1}{2}$. Call $p_1, \ldots, p_k$ the small primes, and the others the big primes. For a fixed $N > 0$ let N_b be the number of $n \leq N$ which are divisible by at least one big prime, and N_s the number of such integers with only small prime divisors. Clearly $N = N_b + N_s$. On the other hand, Erdős shows, by an intricate combinatorial argument, that for a suitable N (in fact 2^{k+2}), $N_b + N_s < N$, which gives the contradiction.

2.3 Proof as Exploration

The above proof is also an example of the third role of proof, that of exploration. Every mathematician knows that when he/she writes out a proof, new insights, ideas, and questions emerge. Moreover, the proof requires techniques which may then be applied to the consideration of new problems. What makes this topic interesting, and somewhat complex, is that there is not always a hard line between explanation and exploration. Often the hallmark of a good proof is that it proves more than the statement of the theorem, as the Erdős proof illustrates.

A fascinating example of proof as exploration is the story of the proof of the alternating sign matrix conjecture, a topic on the boundary of algebra and combinatorics. An alternating sign matrix (ASM) is a square matrix of 0s, 1s, and -1s such that the sum of the entries in each row and each column is 1 and the nonzero entries in each row and each column alternate in sign. These are generalizations of permutation matrices. The ASM conjecture (now the ASM theorem) concerns the number A_n of such $n \times n$ matrices, which is given by $A_n = \prod_{0 \leq j \leq n-1} \frac{(3j+1)!}{(n+j)!}$. (In contrast, there are $n!$ permutation matrices.)

The history of the proof is brilliantly developed in David Bressoud's book *Proofs and Confirmations* [Bressoud 1999]. (The title was inspired by Imre Lakatos' book *Proofs and Refutations* [Lakatos 1976] which in turn was adapted from Karl Popper's *Conjectures and Refutations* [Popper 1963].) Bressoud presents the proof as an exploration, and in fact the chapter containing the proof is entitled "Explorations." He is referring to the development of the proof of the ASM conjecture, which he's presenting the way it developed historically.

Woven into the narrative are classical antecedents of the ASM conjecture, including an algorithm for the evaluation of determinants due to Charles Dodgson (Lewis Carroll), the appearance of many participants (including Mills, Robbins, Rumsey, Stanley, Andrews, and Zeilberger) as well as other results and conjectures. In fact, the ASM conjecture is one of fourteen related conjectures, two of which are still unproved. (One of these was "checked by one of the largest army of reviewers any paper has seen: 88 referees and one computer.")

Bressoud writes that the ASM proof "lay in unexpected territory and revealed a host of new insights and engaging problems." The unexpected territory included plane partitions, symmetric functions, and hypergeometric series. Indeed, it turns out that physicists were interested in ASMs, but they called them six vertex models or square ice.

The strategy of the proof was to try to find a one-to-one correspondence between $n \times n$ ASMs and descending plane partitions with largest part less than or equal to n. (Plane partitions are partitions of integers arranged as a two dimensional array, with certain restrictions. Some of the other conjectures concern generating functions - namely power series whose coefficients count the number of certain plane partitions.)

2.4 Proof as Justification of Definitions

Still another reason for proof, closely connected to teaching, is the justification for mathematical definitions. (I am indebted to my colleague Paul Green for this observation.) For example, one proves that the sum and product of continuous functions is continuous to confirm that the $\varepsilon - \delta$ definition is successful in capturing the intuitive idea of continuity. Similarly, the proof of the intermediate value theorem justifies the definition of the real number system. Yet another example is the use of the fundamental theorem of calculus to show that there is a real valued function whose derivative is e^{-x^2}. What is involved here is the very definition of a function. It demonstrates that a function need not be given by a simple formula, which is something we want to drive home to students. Only a formal proof can guarantee its existence and allow it to be studied.

Proofs also develop and underscore connections between different branches of mathematics, frequently to the benefit of both areas. Well known instances of this phenomenon are combinations of algebra and topology, and of combinatorics and number theory. In "Four Examples," we'll discuss in detail Furstenberg's proof on the Szemeredi theorem, which combines ergodic theory and combinatorial number theory.

2.5 The Dieudonné-Katznelson Encounter

It is frequently asserted that one can fill in the details of an informal argument to obtain a formally correct proof. To quote Bourbaki [1968, p. 8]: "In general [a mathematician] is content to bring the exposition to a point where his experience and mathematical flair tell him that translation into a formal language would be no more than an exercise of patience (though doubtless a very tedious one)." As if in reply, Hersh [1997, p. 52] says "It may be true. It's a matter of faith."

In this connection, let me turn to a personal recollection. In 1971, the distinguished mathematicians Yitzhak Katznelson and Jean Dieudonné visited the University of Maryland for a semester. Katznelson gave a course in ergodic theory, to which Dieudonné was a faithful attendee (as was I). Katznelson's lectures were well organized, although somewhat informal. Dieudonné (who had been a member of Bourbaki) didn't give Katznelson a moment's peace. He kept saying "That is not a proof" or sometimes "That's a nice presentation of the idea—now let's see the proof" and made Katznelson go over the argument until it was accomplished to his (Dieudonné's) satisfaction.

I'm certainly not saying that Dieudonné was more "rigorous" than Katznelson. Katznelson's proofs definitely met the standards of mathematical discourse. There are many acceptable styles of proof. (One might imagine Dieudonné lecturing, with Alonzo Church in the audience, who would say that Dieudonné's arguments were not proofs.)

Dieudonné had a high regard for Katznelson and the course (as he told me) and probably thought that the latter's arguments were essentially correct. But Dieudonné was not playing games. I'm sure he was serious in asserting that Katznelson's arguments fell short of proof, and felt that it was worth the class time for the development of one which was acceptable to him.

2.6 The Jaffe-Quinn Article

An extremely interesting discussion of various issues concerning proof was initiated by an article in the *Bulletin of the American Mathematical Society* by Arthur Jaffe and Frank Quinn [1993], and the responses it generated [Responses 1994]. The article (henceforth referred to as JQ) is entitled "Theoretical mathematics: towards a cultural synthesis of mathematics and theoretical physics." JQ use the term "theoretical mathematics" for "speculative and intuitive work" (this terminology was much criticized by a number of the respondents) and "rigorous mathematics" for "proof oriented work." While they agree that mathematics is "nearly characterized by the use of rigorous proofs" (which they unequivocally endorse) they call attention to "a trend towards basing mathematics on intuitive reasoning without proof" and say that this "may be the beginning of fundamental changes in the way mathematics is organized."

JQ contrast mathematics with physics. In the latter there is a "division of labor" between experimenters and theoreticians. But "the mathematical community has not undergone a bifurcation into theoretical and rigorous branches."

There is at least an implication by JQ that such a "bifurcation" would be desirable. But the lack of it is not accidental, and I doubt that it can be created by fiat. Of course there always has been a speculative and intuitive component to mathematics (and JQ correctly point to this as one of mathematics' "success stories") but I don't think there can be a division of mathematicians into two kinds, as there is in physics. That is, in general a mathematician's work is both intuitive *and* rigorous. Certainly there are individuals—Mandelbrot (one of the respondents) and Feigenbaum come to mind—whose main activity is "theoretical," but it's doubtful that there will be an entire community of such.

A year later the *Bulletin* printed a number of responses to JQ (by pure and applied mathematicians, physicists, and a historian of mathematics), as well as a separate article by Bill Thurston. These were in turn followed by a response by JQ.

While some of the responders are in substantial agreement with JQ, there are attacks from both the "right" and the "left." Mac Lane felt that physics is not a good model for mathematics. (The quote from Mac Lane in the section "The Roles of Proof," above, is part of his response.) Moe Hirsch (presumably tongue in cheek) suggests that "published mathematics . . . like good wine, should carry a date. If after ten years no errors have been found the theorem will be generally accepted" and that one should "attach a label to each proof, e.g., computer aided, mass collaboration, formal, informal, constructive, fuzzy, etc."

A particularly negative response was by Benoit Mandelbrot. He finds JQ "appalling" and refers to rigorous mathematicians as "Charles" mathematicians (since the AMS office in Providence is on Charles Street). He characterizes mathematical rigor as "besides the point and usually distracting, even where possible."

Richard Palais, the editor of the *Bulletin*, wrote that ("with mixed feelings") the *Bulletin* would no longer publish "controversial" articles. (Such would be restricted to the *Notices of the American Mathematical Society*.) One wonders about the subtext of this decision.

3 Computers and Proof

There is no question that computers are having a profound impact on mathematical practice. Perhaps their main role has been in experimentation, production of pictures, data, and the generation of conjectures. But computers have been used in some controversial proofs.

The relation between computers and proof is quite complex, and is still being sorted out. This paper will consider only a few such cases. It is interesting that Rota, in a passage following the quotation cited in the section "The Roles of Proof," above, says that it is *because* of computers that proof is "more indispensable than ever" (since "conjectures in number theory may fail for integers . . . beyond the reach of . . . computers"). There are some mathematicians, notably Paul Halmos and Pierre Deligne, who completely reject the use of computers. For example, Deligne has written "I don't believe in a proof done by a computer . . . I believe in a proof if I understand it." [Szpiro 2003, p. 21] In the same spirit, Eugene Wigner is reported to have said [Robertson 2003, p. 80] "It's nice to know that the computer understands the problem. But I would like to understand it, too." On the other hand, Thomas Hales says "I now feel that computer proofs are vital to the progress of mathematics." [Szpiro 2003, p. 212]

I take an intermediate point of view. Regardless of anyone's feelings (even Deligne's), one cannot wish away the use of computers in proofs. Mathematicians will use them if they find them necessary, or even convenient, and it's necessary to come to terms with this phenomenon. On the other hand, it's somewhat disingenuous to say that there is no difference between a calculation done by a computer and one done "by hand."

The issue is not whether one should "believe" a proof making use of a computer. Indeed, it may well be the case that a computer calculation is more reliable than a traditional one, especially if the latter is very long (witness the competing attempts at proving the Kepler conjecture, discussed below). Some of the same processes as in traditional proofs, for example modifications of the original argument, and repeated scrutiny, occur with computer proofs, and confirm the truth of the claimed assertion.

Moreover, there are certain proofs which just couldn't be accomplished without a computer. One such is the much discussed proof of the four color theorem, by Appel and Haken. The problem was reduced to several thousand cases, which were then checked by the computer.

The point is that it is necessary to recognize that there are tradeoffs involved here, namely the achievement of results versus the understanding of the reasons for their proofs. Even a rote computation in a traditional proof involves a certain amount of thinking. In the case of replication of a computer argument we cannot determine easily what hidden assumptions or errors lie in the shared bits of coding or hardware. At some point in the proof, a result is true because the computer "said so."

With regard to computers and proof, the story of the Kepler conjecture on sphere packing is particularly striking. (In my opinion, it is an order of magnitude more interesting than the four color theorem, although the latter was the first well known problem to make use of the computer for its solution).

The conjecture is that the densest way to pack spheres is the hexagonal close (or "greengrocers") packing. This is a four hundred year old problem, the oldest problem in discrete geometry, which was also part of Hilbert's 18th problem. There was a disputed proof (by Hsiang), and then a very long, computer assisted proof (by Hales), which is apparently correct. And the latter has led to conjectures and proofs of new results. All of this is recounted in detail in the excellent book *Kepler's Conjecture* by George C. Szpiro [2003].

A proposed proof, by Wu-Yi Hsiang [1993], was actually published. This proof made no use of the computer, just tools from (relatively) elementary geometry and calculus. The consensus of the mathematical community is that the attempted proof is incorrect, although Hsiang still stands by it. The proof that is now generally accepted is by Thomas Hales, with significant help from his student Samuel Ferguson. It consists of six papers, as well as a computer program. It was submitted to the *Annals of Mathematics* (in fact it was solicited by the *Annals*) and a team of 12 referees worked on it for four years. They returned a report saying that they were unable to completely certify the proof, although they were 99 percent certain of it.

In fact, the *Annals* has published Hales' proof [Hales 2005], although not the computer code on which it was based. The original plan was to publish it with a disclaimer, but after Hales reorganized it, it appeared as a single (more than one hundred page) paper, without a disclaimer. On the first page, Hales writes, "Here we describe the top-level outline of the proof and give sources of details of the proof. The latter are to appear as several papers in *Discrete and Computational Geometry*."

The Szpiro book has a chapter entitled "But is it really a proof?" There does seem to be a strong consensus that the Kepler conjecture is now proved—that it is "certified." There is considerably less agreement as to whether it meets the criterion of "explanation." For example, the mathematician and science writer Ian Stewart likens Hales' proof to a telephone directory, in contrast to Wiles' proof of Fermat's last theorem, which he compares to "War and Peace."

In this case, how are we to decide if the "telephone book" nature of Hales' proof is inherent to the problem?

This type of thing is unprecedented in mathematics. Regardless of one's feelings about the use of computers in proof, it must be recognized that Hales' work is a major scientific achievement. The story is not over; although the proof has appeared in print, there will very likely be simplifications that will really embed the result into mathematics.

The proofs of the four color theorem and the Kepler conjecture definitely fall within the traditional framework of proof as a sequence of deductions, although the computer plays an essential role. But there is another trend which in fact challenges the accepted dichotomy between a proof and an argument which falls short of proof. This is not concerned with the computer as an aid to proof, but rather envisions computer calculations as replacing proof.

One of the most provocative challenges to traditional proof was put forth by Doron Zeilberger in an article "Theorems for a price: tomorrow's semi-rigorous mathematical culture." [Zeilberger 1993] (As we'll see, "for a price" is meant literally.) The tone is set by Zeilberger's much quoted (and by now notorious) statement that in the future "rigorous old style mathematicians . . . may be viewed by mainstream mathematicians as a fringe sect of harmless eccentrics." He continues: "The computer has already started doing to mathematics what the telescope and microscope did to astronomy and biology. . . . In the future mathematicians will not care about absolute certainty, since there will be so many exciting new facts to discover." After presenting a number of identities which were proved by, or with the aid of, a computer, he envisions an abstract of a paper (c. 2100); "We show in a certain precise sense that the Goldbach conjecture is true with a probability larger than 0.9999 and that its complete truth could be determined with a budget of $10 billion." (Perhaps intentionally, there is no explanation of this assertion by the 2100-era mathematician.)

I should mention that Zeilberger is an outstanding mathematician, and in fact was one of the participants in the solution of the alternating sign matrix conjecture. But on this question I think he is quite wrongheaded.

Zeilberger's article was reprinted in the *Mathematical Intelligencer*, where it is followed by a response from his friend and collaborator George Andrews [1994]. Andrews' article is entitled (in part) "You've got to be kidding." He challenges Zeilberger's evaluation of the role of the computer in the discovery and proof of the various identities, and says moreover that Zeilberger "ignores the insight provided by proof" and "has produced exactly no evidence that his Brave New World is on the way."

As for Zeilberger's assertion that important theorems can be proved "for a price"—I don't believe it. Mathematics just doesn't work that way. Although mathematicians are no more immune to the lure of money than anyone else, one can't imagine a "crash program" to prove the Riemann hypothesis or the twin prime conjecture. It's true that there is now a well publicized monetary prize for such proofs, but there is no reason to think that the proofs will be attained any earlier on that account.

Another proponent of this trend is the geologist Douglas Robertson [2003]. Robertson's point of view is similar to Zeilberger's (he might be termed "Zeilberger lite"). Interestingly, he is extremely frank and explicit about what may be lost by this process. "Just as astronomers had to accept the idea that the telescope vastly extends the reach of the naked eye, mathematicians will have to accept the idea that the computer similarly extends the reach of the human mind." [Robertson 2003, p. 81] He also says that the understanding of the reasons behind such a computer proof "may not be attainable." Robertson asserts that computers will throw light on whether π is a normal number.[1] (This is very doubtful, in my opinion. No amount of computer calculation can settle this question.)

4 Four Examples

The examples which I'll discuss at some length are from topological dynamics and ergodic theory, areas of which I have some knowledge.

[1] A number is said to be normal (say to base 10) for which every finite sequence in the decimal expansion occurs with the "right" limiting frequency. For example, the occurrence of 57 has limiting frequency .01. Normal numbers have full Lebesgue measure, but are of first category.

4.1 The Birkhoff Ergodic Theorem

Even an absurdly naive idea can lead to a valid proof. Recall the statement of G.D. Birkhoff's (pointwise) ergodic theorem: Let T be a measure preserving transformation on a probability space X and let f be an integrable function on X. Then $\lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n} f(T^i x)$ exists for almost all x. (T^idenotes the i fold composition of the transformation T.)

In his book *Lectures on Ergodic Theory* [Halmos 1956], Halmos, after proving and obtaining some consequences of the ergodic theorem, concludes a chapter with what he calls an alternative "proof."

If f is a non-negative function on the positive integers, write

$$\int f(n)dn = \lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} f(i)$$

whenever the limit exists, and call such functions integrable. If T is a measure preserving transformation on a space X and f is an integrable function on X, then

$$\iint |f(T^n x)|dn\,dx = \iint |f(T^n x)|dx\,dn = \iint |f(x)|dx\,dn = \int |f(x)|dx < \infty.$$

Hence by "Fubini's theorem"(!) $f(T^n x)$ is an integrable function of its two arguments, and therefore, for almost every fixed x, it is an integrable function of n. That is,

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} f(T^i x)$$

exists for almost all x.

(This "proof" would work if there were a probability measure on the integers which assigned equal measure to each integer.)

One might think that this was just a joke on the part of Halmos, but he then asks, "Can any of this nonsense be made meaningful?" In fact, more than thirty years later, Ornstein obtained a proof based on this idea. This is the most conceptual proof of the ergodic theorem. The proof is by a delicate approximation argument. Essentially, one considers the product of the measure space with a finite space. This is a good example of the role of explanation in proof. Other proofs, some of which are shorter, depend on a trick. With this proof one really sees what is happening. The ideas of the proof led to theorems (by Ornstein and Weiss) on the actions of locally compact amenable groups.

4.2 Topological Entropy

Lakatos characterizes mathematical knowledge as proceeding by a sequence of proofs and refutations, and indeed sometimes erroneous "proofs" lead to refinements and clarifications. A case in point occurred in the early work on entropy, which is an important numerical invariant for dynamical systems.

Actually, there are two kinds of entropy. Measure theoretical entropy (defined by Kolmogorov and Sinai around 1958) applies to measure preserving transformations of probability spaces, and is defined in terms of measurable partitions. Topological entropy, defined several years later by Adler, Konheim, and McAndrew [Adler *et al.* 1965] concerns continuous self maps of compact metric spaces, and is defined in terms of open sets. In each case, the entropy is an extended non-negative real number. That is, if T is the transformation defined on the (probability or metric) space X, then the entropy $h(T)$ satisfies $0 \leq h(T) \leq \infty$. There are interesting connections between the two notions of entropy.

For measure theoretic entropy, there is a product theorem. If T_1 and T_2 are measure preserving transformations of X_1 and X_2 respectively, and $T_1 \times T_2$ is the product transformation, then $h(T_1 \times T_2) = h(T_1) + h(T_2)$. Moreover, the proof follows almost immediately from the definition.

When Adler, et al. introduced topological entropy, their paper included what they presented as a proof of the corresponding theorem for continuous transformations, which closely mimics the measure theoretic proof. This "proof," which was published, was erroneous—it slipped by a careless referee.[2]

This error (which was discovered by Kakutani) inspired an alternative (equivalent) definition of topological entropy by Bowen, which is in some ways more useful than the original definition. In particular, a correct proof of the product theorem can be given using this definition (although as a matter of fact Bowen's first proof was also incorrect).

4.3 The Szemeredi Theorem

Many important theorems have more than one proof. On an elementary level, there are several proofs of the infinity of primes, and several hundred proofs of the quadratic residue theorem (ten by Gauss). Also, as was mentioned above, there are a number of proofs of the ergodic theorem.

Of course, in terms of validation of a result, one correct proof is sufficient. If there is any question, the existence of multiple proofs provides some confirmation that a proposed result is correct. In this respect, mathematics is like an experimental science. An alternative proof is something like a replication of an experiment.

But I think what is even more important is that a new proof frequently connects with other branches of mathematics. As Michael Atiyah [2005] says "different proofs have different strengths and weaknesses, and they generalize in different directions—they are not just repetitions of each other."

A striking example is the Szemeredi theorem, which says that every set of integers of positive upper density contains arbitrarily long arithmetic progressions. Szemeredi's original proof was combinatorial and extremely long. Hillel Furstenberg gave another proof, which was accomplished by translating the problem to ergodic theory.

The main ergodic theoretic lemma is:

Let $T_1, \ldots, T_k$ be commuting measure preserving transformations of a probability space (X,μ), and let A be a set of positive measure. Then there is a positive integer n such that $\mu(A \cap T_1^{-n} A \cap \cdots \cap T_k^{-n} A) > 0$. To see how this measure theoretic result implies Szemeredi's theorem, we consider $\Omega = \{0, 1\}^Z$, the space of doubly infinite sequences of zeroes and ones, provided with the product topology. The shift transformation T on Ω is defined by $T\omega(n) = \omega(n+1)$.

Now let S be a subset of the integers of positive upper density, and let 1_S be the indicator function of S (that is $1_S(n) = 1$ if $n \in S$ and 0 otherwise); 1_S is a point of Ω. Let X be the orbit closure of 1_S under T, and let $A = \{\omega \in X \mid \omega(0) = 1\}$. It can be shown that there is a measure μ on X which is invariant under the shift for which $\mu(A) > 0$. (This fact depends on, and in fact is equivalent with the assumption that S has positive upper density.) Now apply the above lemma to the commuting transformations $T, T^2, \ldots, T^k$. It follows that there is a point $\omega \in X$ for which

[2] I was the careless referee.

$T^{jn}(\omega) \in A$ for some n and $j = 1,2,\ldots,k$ from which one easily deduces that for some h, one has $h, h+n, \ldots, h+kn \in S$.

Byproducts of the proof include a general structure theorem for ergodic transformations, which in turn inspired an analogous structure theorem (by Veech) in topological dynamics, as well as "multidimensional" Szemeredi theorems. Furstenberg's proof initiated a fruitful connection between ergodic theory and combinatorial number theory. The ideas introduced played a role in the spectacular recent work of Tau and Green on the existence of arbitrarily long arithmetic progressions in the primes [Green/Tau to appear].

4.4 A Fixed Point Theorem

Standards of proof vary over time.[3] For example, it's well known that Euclid's proofs were incomplete (although apparently all of his theorems are correct). Also, the proofs of the Italian algebraic geometers of the early part of the last century are now found wanting.

An interesting more recent case is provided by a paper by Morton Brown and Walter Neumann [1977], which is related to two papers of G.D. Birkhoff, [1913] and [1925]. Birkhoff claimed to prove a conjecture of Poincaré ("Poincaré's last geometric theorem"), which asserted the existence of two fixed points for an area preserving homeomorphism of an annulus which rotates the boundary circles in opposite directions. Over the years there were questions as to whether Birkhoff's proof was correct. The paper of Brown and Neumann presents a proof which the authors generously say is essentially the same as Birkhoff's.

Be that as it may, the language of Brown and Neumann is quite different from Birkhoff's, reflecting the development of topology in the intervening years. Some of Birkhoff's statements lacked precision. For example a curve is defined to be the boundary of an open set. It isn't even clear what is meant by rotating the boundary curves in opposite directions. (A clockwise rotation of one degree can be regarded as a counterclockwise rotation of 359 degrees.) In the Brown-Neumann paper, this is made precise by passing to the universal covering space, a notion which was probably not known to Birkhoff. Another tool is the homotopy lifting property, which also was probably not known explicitly to Birkhoff.

One expects that fifty years from now, some of the proofs of mathematicians of 2006 will be thought to be in need of correction or modification.

While there are differences among the four proofs just discussed, a common thread is what might be called reinforcement (a new proof, or a correction, or a reworking of an earlier proof). There seems to be no doubt that the theorems in question have been proved.

We are confronted with a different situation with certain very long proofs. We conclude with a brief discussion of two current (possible) proofs of important results, on which the jury is still out.

One is the classification of finite simple groups, organized by the late Daniel Gorenstein, of which there is some doubt whether it has actually been accomplished. (Moe Hirsch, in his response to JQ irreverently asks "Who's in charge here, anyway?")

The other is the (apparent) proof by Perelman of Thurston's geometrization conjecture (which implies the Poincaré conjecture). John Morgan, in a survey article [Morgan 2005], writes

[3] A fascinating discussion of differing standards of proof over time is presented in [Kleiner/Movshovitz-Hadar 1997].

"The mathematical community is still trying to digest his argument and ascertain whether it is indeed . . . complete and correct."

In the latter case, the expectation is that a proof will in fact emerge. In spite of certain differences (in particular, the computer plays no role in Perelman's arguments) something like the "Kepler process" is occurring. As was discussed above, following a lengthy and elaborate process, Hales' proof is now generally accepted, and it's quite likely the same will hold for Perelman's.[4]

On the other hand, opinion is sharply divided in regard to the classification of finite simple groups, and it's anyone's guess as to how it will finally turn out.

Acknowledgements I would like to thank the following: Ethan Akin, Stuart Antman, Ken Berg, Mort Brown, Bill Byers, Marshall Cohen, Keith Devlin, Hillel Furstenberg, Bonnie Gold, Judy Green, Paul Green, Denny Gulick, Frances Gulick, Brian Hayes, Karl Hofmann, Michael Keane, Jonathan King, Barbara Meeker, Jonathan Rosenberg, Doris Schattschneider, Tom Ward.

References

[Adler *et al.* 1965] Roy Adler, A.G. Konheim, and M.H. McAndrew, "Topological entropy," *Trans. Amer. Math. Soc.* 114 (1965), pp. 309 – 319.

[Andrews 1994] George G. Andrews, "The death of proof? Semi-rigorous mathematics? You've got to be kidding!" *The Mathematical Intelligencer* 16 (1994), pp. 16–18.

[Atiyah 2005] "Interview with Michael Atiyah and Isadore Singer," *Notices Amer. Math. Soc.* 52 (2005), pp. 225–233.

[Aigner/Ziegler 1999] Martin Aigner and Gunter M. Ziegler, *Proofs from THE BOOK*, Springer, 1999.

[Bass 2003] Hyman Bass, "The Carnegie initiative on the doctorate: the case of mathematics," *Notices Amer. Math. Soc.* 50 [2003], pp. 767–776.

[Bing 1948] RH Bing, "A homogeneous plane continuum," *Duke Mathematics Journal* 15 (1948), pp. 729–742.

[Birkhoff 1913] G.D. Birkhoff, "Proof of Poincaré's last geometric theorem," *Trans. Amer. Math Soc.* 14 (1913), pp. 14–22.

[Birkhoff 1925] ——, "An extension of Poincaré's last geometric theorem," *Acta Mathematica* 47 (1925), pp. 297–311.

[Biss 2004] Daniel Biss, "The elephant in the internet," *Notices Amer. Math. Soc.* 51 (2004), pp. 1217–1219.

[Bourbaki 1968] Nicholas Bourbaki, *Elements of Mathematics, Theory of Sets*, Addison-Wesley, 1968.

[Bressoud 1999] David Bressoud, *Proofs and Confirmations, The Story of the Alternate Sign Matrix Conjecture*, Cambridge University Press, 1999.

[Brown/Neumann 1977] Morton Brown and Walter Neumann, "Proof of the Poincaré -Birkhoff fixed point theorem," *Michigan Math. Jour.* 24 (1977), pp. 21–31.

[4] The apparent confirmation of Perelman's proof has come too late for detailed consideration in this article. Perelman has posted his papers on his website, but has refused to submit them for publication. As we have noted, this is contrary to accepted scientific practice. In any case, there has been extensive discussion of the proof, including articles in the *New York Times* (August 15, 2006) and *The New Yorker* (August 25, 2006).

[Choquet 1944] G. Choquet, "Prolongement d'homéomorphes," *Comptes Rendus* 219 (1944), pp. 542–544.

[Davis/Hersh 1981] Philip Davis and Reuben Hersh, *The Mathematical Experience*, Houghton Mifflin, 1981.

[Ernest 1998] Paul Ernest, *Social Constructivisn as a Philosophy of Mathematics*, State University of New York 1998.

[Franks 1989] John Franks, "Comments on the responses to my review of *Chaos*," *Mathematical Intelligencer* 11 (1989), pp. 12–13.

[Furstenberg 1981] H. Furstenberg, *Recurrence in Ergodic Theory and Combinatorial Number Theory*, Princeton University Press, 1981.

[Gale 1990] David Gale, "Proof as explanation," *Mathematical Intelligencer* 12 (1990), p. 4.

[Green/Tau to appear] Ben Green and Terry Tau, "The primes contain arbitrarily long arithmetic progressions," *Annals of Mathematics*, to appear.

[Hales 2005] Thomas Hales, "A proof of the Kepler conjecture," *Annals of Mathematics* 162 (2005), pp. 1063–1183.

[Hales www.] ——, Flyspeck project fact sheet (www.math.pitt.edu/~thales/flyspeck/).

[Halmos 1956] Paul R. Halmos, *Lectures on Ergodic Theory*, Mathematical Society of Japan, 1956.

[Hersh 1997] Reuben Hersh, *What is Mathematics, Really?* Oxford University Press 1997.

[Hsiang 1993] Wu-Yi Hsiang, "On the sphere packing problem and the proof of Kepler's conjecture," *International Journal of Mathematics* 4 (1993), pp. 739–781.

[Jaffe/Quinn 1993] Arthur Jaffe and Frank Quinn, "Theoretical mathematics: towards a cultural synthesis of mathematics and theoretical physics," *Bull. Amer. Math. Soc.* 29 (1993), pp. 1–13.

[Kitcher 1984] Philip Kitcher, *The Nature of Mathematical Knowledge*, Oxford University Press, 1984.

[Kleiner/Movshovitz-Hadar 1997] Israel Kleiner and Nitsa Movshovitz-Hadar, "Proof: a many-splendored thing," *Mathematical Intelligencer* 19 (1997), pp. 16–26.

[Lakatos 1976] Imre Lakatos, *Proofs and Refutations: The Logic of Mathematical Discovery*, Cambridge University Press, 1976.

[Lewontin 2005] Richard Lewontin, "On fraud in science: an exchange," *New York Review of Books*, February 10, 2005, pp. 46–48.

[Morgan 2005] John Morgan, "Recent progress on the Poincaré conjecture and the classification of 3-manifolds," *Bull. Amer. Math. Soc.* 42 (2005), pp. 57–78.

[Popper 1963] Karl Popper, *Conjectures and Refutations: The Growth of Scientific Knowledge*, Routledge, 1963.

[Responses 1994] "Responses to 'Theoretical mathematics,'" *Bull. Amer. Math Soc.* 30 (1994), pp. 178–207.

[Robertson 2003] Douglas Robertson *Phase Change, the Computer Revolution in Science and Mathematics*, Oxford University Press, 2003.

[Ross 1998] Kenneth A. Ross, "The place of algorithms and proofs in school mathematics," *American Mathematical Monthly* 105 (1998), pp. 252–255.

[Rota 1993] Gian-Carlo Rota, "The concept of mathematical truth," pp. 91–96 in Alvin White, ed., *Essays in Humanistic Mathematics*, Mathematical Association of America, 1993.

[Rota 1996] ——, "The phenomenology of mathematical proof," pp. 134– 150 in Gian-Carlo Rota and Fabrizio Palombi, *Indiscrete Thoughts*, Birkhauser, 1996.

[Szpiro 2003] George G. Szpiro, *Kepler's Conjecture*, John Wiley, 2003.

[Waraskewich 1937] Z. Waraskewich, "Sur les courbes planes topologiquement homogènes," *Comptes Rendus* 204 (1937), pp. 1388–1390.

[Weinberg 2001] Steven Weinberg, "Can science explain everything? Anything?" *New York Review of Books*, May 31, 2001, pp. 47–50.

[Yandell 2001] Benjamin Yandell, *The Honors Class: Hilbert's Problems and their Solvers*, A.K. Peters, 2001.

[Zeilberger 1993] Doron Zeilberger, "Theorems for a price: tomorrow's semi-rigorous mathematical culture," *Notices Amer. Math. Soc.* 40 (1993), pp. 978–981.

II

Social Constructivist Views of Mathematics

Two completely new philosophies of mathematics have been developed since 1950: structuralism and social constructivism. Structuralism is the view that mathematics is the science of structures, or patterns. That view is discussed in several of the chapters in section 3. Social constructivism has been developed primarily by mathematicians, although one can trace its origins to some discussion by philosophers such as Lakatos. Social constructivism is the view that mathematics is constructed by the community of mathematicians. In one sense, this is so obviously true that there is no need to discuss it further. Certainly, human *knowledge* of mathematics *is* developed by the community of mathematicians. However, as we discover mathematical facts, it *feels* to most of us as if there is an objective reality out there, within which these facts are either true or false. It certainly does not seem that the bunch of us can just one day decide, "the Riemann hypothesis is true," and it will be so. On the other hand, when a new mathematical *concept* is introduced and developed, things are less clear. Is there some external "natural" concept that we're grasping for? Or are we just making it up, albeit with some restrictions related to the questions we are developing it to investigate? The less extreme versions of social constructivism, represented in this volume, suggest that, once the community has developed a mathematical concept, the facts about this concept are indeed objective. However, there are philosophical issues with this viewpoint, and these are also discussed in this section.

The first two chapters of this section were written by mathematicians who have been outspoken proponents of social constructivism, and the third is by a philosopher who has been working to formulate social constructivism carefully enough for criticism by the community of philosophers of mathematics.

4

When Is a Problem Solved?[1]

Philip J. Davis
Division of Applied Mathematics
Brown University

From the Editors

The question Philip Davis asks in his chapter, "When is a problem solved?" seems like a natural one to ask, but we have never read a discussion of this elsewhere. It is a good example of why it is important for some people who actually do mathematics to contribute to the philosophy of mathematics. There are questions of interest to mathematicians that do not occur to philosophers, who are motivated largely by the types of questions that occur in other areas of philosophy. This question might never occur to philosophers, because it is really only in mathematics that we appear *to get final answers to our questions.*

Philip Davis is a Professor Emeritus of Applied Mathematics at Brown University (www.dam.brown.edu/people/facultypage.davis.html). He came to Brown after serving as Chief for Numerical Analysis at the National Bureau of Standards in Washington, D.C. for five years. His fields of research included numerical analysis and approximation theory, in which he wrote many papers and several books, including Interpolation and Approximation *(1963),* Numerical Integration *(with Philip Rabinowitz, 1967), and* The Schwarz Function *(1974) and* Circulant Matrices *(1979). He is a prize winning expositor of mathematics, who received the Chauvenet Prize of the Mathematical Association of America in 1963 for "An Historical Profile of the Gamma Function." Professor Davis has also received the Laster Ford Award in 1982 for "Are there Considences in Mathematics?" and the George Polya Award in 1986 for "What Do I know? A study of Mathmatical Self-Awareness." In 1997, he won the Communications Award of the Joint Policy Board for the Mathematical Science. His books written jointly with Reuben*

[1] This article was initially submitted to this volume. However, due to the long time between initial solicitation of articles and the appearance of this volume, this article first appeared in 2006 in a collection of my articles, *Mathematics and Common Sense: A Case of Creative Tension*. It is reprinted here with permission of the publisher, AK Peters.

Hersh, The Mathematical Experience *(1980) and* Descartes' Dream *(1986), explore certain questions in the philosophy of mathematics, and the role of mathematics in society.* Mathematics and Common Sense: A Case of Creative Tension, *which appeared in 2006, contains a version of his chapter, among other philosophical articles. Readers of this volume will also be interested in his article, "When Mathematics Says No" in* No Way: The Nature of the Impossible, *which he edited with David Park (1987).*

A poem is never finished, it is only abandoned.—Paul Valéry

1 Introduction

I recently spent three days participating in MathPath, a summer math camp for very bright students aged c. 12–14 (see www.mathpath.org). One day I asked the students to pass in to me a question that was a bit conceptual or philosophical. Out of the large variety of responses, one question struck me as both profound and remarkable in that sophisticated interpretations were possible:

Elizabeth Roberts: ***How do we know when a problem is solved?***

My first reaction on reading this question—which was pencilled on a sheet of notebook paper—was "mathematical problems are never solved." Due to my limited stay at the camp, I didn't have the opportunity to ask the student what exactly she meant and so her question went unanswered at the time. I told the camp faculty—all professional mathematicians—my gut reaction. I added that my answer was ***not*** appropriate for the present age group and hoped that the faculty would take up the question after I'd left. I also told the faculty that the question inspired me to write an article. Here it is.

2 A Bit of Philosophy

Some problems are solved. A baker knows when a loaf of bread is done.[2] Yogi Berra said: "It's not over till it's over." Which implies that a baseball game gets over. But when one thinks of the problems that confront humanity—personal, medical, sociological, economic, military—problems that seem **never** to be solved, it is easy to conclude that to be truly alive is to be perpetually racked by problems.

Example: When should clinical trials for new medical procedures be terminated? This question is currently on the front pages of newspapers and is a matter of litigation and the confrontation of statisticians involved in the jurimetrics. [Finkelstein/Levin 2004]

Thus, we are concerned here with a fundamental question that can be viewed as residing at the heart of human existence itself. How can we be sure that we have solved a problem? More than this, how can we be sure we have formulated a proper question? We can't, because problems, questions and solutions are not static entities. On the contrary, the creation, formulation and solution of problems change throughout history, throughout own lifetime and throughout our

[2] In an amusing e-letter, Yvon Maday, a Parisian applied mathematician, pointed out to me ambiguities in the baking process.

readings and re-readings of texts. That is to say, meaning is dynamic and ongoing and there is no finality in the creation, formulation and solution to problems, despite our constant efforts to create order in the world. Our ability to create changes in meaning is great and hence our problems and our solutions change. We frequently settle for provisional, "good enough" solutions—often described as "band aid solutions." [O'Halloran 2005]

3 What Might Elizabeth Have Meant?

One might think that in the case of mathematics—that supposedly clean-cut, logical, but limited intellectual area—the situation would be otherwise. One might think that when a mathematical problem arises, then after a while (it may be a very long while) the problem gets solved. But think again; what takes place can be very complex.

The set of possible responses to the question under discussion spans the whole of mathematical methodology, history, and philosophy. Though responses are implicit everywhere in the mathematical literature, I believe that the question as framed here puts a slightly different slant on this material. I don't recall seeing it treated head on.

The question: **How do we know when a problem is solved?** can be approached at a variety of levels. The lay public tends to think that mathematics is an area where there is one and only one answer to a problem. Approached from the point of view of a school teacher, the teacher, relying on habits or traditions, and considering the age of the pupils, knows when a pupil has solved a problem. It is a matter of common sense. (I am not thinking here of multiple choice questions graded by machine.)

Approached from the point of view of the individual or the group that makes up problems either for daily work, tests, or contests, I would suppose that the act of making up the problem already implies a more or less definite notion of what the answer is. The examiner will think the problem is solved if he gets the answer he had in mind or possibly a variant that conforms to certain unconsciously maintained criteria.

One answer, appropriate to students starting algebra, might be "you plug your solution back into the equation and see if it checks." The set of possible responses that lie between this simplistic response and my seemingly dismissive "mathematical problems are never solved," spans the whole of mathematical methodology, history, and philosophy. Though responses to the question under discussion are implicit everywhere in the mathematical literature, I believe that the question as framed puts a slightly different slant on this material. I don't recall seeing it treated head on.

What did the student mean by her question? I can only guess. Perhaps she meant: "How can I tell whether my answer is correct." Well, what methods or practices of validation are available at ages 12–14? Yes, you can plug the answer back into the equation and see if it checks. But this kind of check is not available for most problems—as, for example, what and where do you plug in when asked to add a column of numbers? If you care to employ them, processes such as "casting out nines" (taught in elementary school years and years ago) or estimating the sum provide partial checks for addition.

You can "check your work" by doing the problem over again in perhaps a simpler or a more clever way and then compare. You may, in some cases, put the problem or part of it on a computer. You can ask your friend what her answer is and compare. You can look in the back of the book and see whether you get the book's answer. If the problem is a "word problem," you can ask whether your answer makes sense in the "real world." An answer of minus seven and a half dappled cows is evidence of an error somewhere.

Perhaps the student, having learned that $\surd 2$ is irrational, will wonder whether or why $\surd 2 = 1.41421356237\ldots$ constitutes an answer. From a certain point of view, $\surd 2$ can never have a completed answer. Does one have to elaborate the meaning of the three dots . . . and trot out the theory of the set of real numbers to accept this as an answer?

Iterative computations that theoretically "converge at infinity" are frequent. They must be terminated—abandoned—and an "answer" outputted. I know at least thirteen different termination criteria that are employed. It would be useful to have a full taxonomic study of such criteria, but I am not aware of such a study.[3]

Perhaps the student, having heard from the camp faculty (or from reading newspapers) that some mathematical problems have taken centuries before they were resolved, was asking me how long she should spend on a problem before abandoning it. We all abandon problems. Life calls us to other things that must get done.

4 *Mathematical Argumentation as a Mixture of Materials*

Here is a final conjecture as to what might have been in the student's mind in asking the question. It is a **very unlikely** conjecture, but it expresses a feeling that I occasionally have after reading through mathematical material.

What is the source of one's confidence that the informal, patched together mixture of verbal argumentation, symbol manipulation, computation and the use of visuals, whether in the published literature or of one's own devising, all click together properly as presented, and result in the confident assertion: "Yes, that certainly solves the problem!"

Let me elaborate. Consider the processes and techniques used in solving mathematical problems. The mélange of materials involved has been well described by mathematical semioticist Kay O'Halloran who studies the relationship between mathematical ideas and the symbols with which these ideas are expressed.

> "Mathematical discourse succeeds through the interwoven grammars of language, mathematical symbolism and visual images, which means that shifts may be made seamlessly across these three resources. Each semiotic resource has a particular contribution or function within mathematical discourse. Language is used to introduce, contextualize, and describe the mathematics problem. The next step is typically the visualization of the problem in diagrammatic form. Finally, the problem is solved using mathematical symbolism through a variety of approaches which include the recognition of patterns, the use of analogy, an examination of different cases, working backwards from a solution to arrive at the original data, establishing sub-goals for complex problems, indirect reasoning in the form of proof by contradiction, mathematical induction and mathematical deduction using previously established results." [O'Halloran 2005]

[3] Each special problem may develop its own special termination criteria. See,e.g., [Ehrich 2001].

Behind the understanding of and expertise with symbolisms, there are cognitive capacities that act to create and glue together the mathematical discourse. Lakoff & Núñez give a list required for doing simple arithmetic. They are (with these authors' elaborations omitted):

> "grouping capacity, ordering capacity, pairing capacity, memory capacity, exhaustion detection capacity, cardinal number assignment, independent-order capacity, combinatorial-grouping capacity, symbolizing capacity, metaphorizing capacity, conceptual-blending capacity." [Lakoff/Núñez 2000]

Just as logicians have wondered whether further axioms are necessary for mathematics, I wonder whether further mental capacities than those above are required to do mathematics that is more complex than simple arithmetic. I wonder whether as mathematics progresses, and as it adds new proofs and develops new theories, we are now in the possession of additional mental capacities in virtue of the work of the brilliant mathematicians of the past. I wonder also whether semantics, semiotics, and cognitive science, taken together, are adequate to explain the occurrence of the miraculous epiphany "Yes. That's it. The problem is now solved." Psychological studies and autobiographical material have not yet uncovered all the ingredients that make up the "aha" moment.

5 From a Mathematician's Perspective

I am now lead to imagine that the question ***How do we know when a problem is solved?*** has been put to a professional. There is no universal answer to this question. It depends on the situation at hand. The typical answers for validation just given to young math students, carry into the professional domain. Examples: product barcodes have check digits that employ modular arithmetic. When, in the first generation of computers, I computed the Gaussian weights and abscissas for approximate integration to 30 D, I plugged back to verify my output. The modes of validating a long and involved computation may involve reworking the problem with a different algorithm, with different software on a different computer and then comparing.[4]

But there is much, much more that has to be said. At the very outset, one might ask: does the problem, as stated, make sense or does it need reformulation? There are ill-posed problems, in either the technical sense or a broader sense. There are well-posed problems, weakly-well-posed problems, etc. One might also ask—but is rarely able to ask at the outset—does the problem have a solution? From the simplest problems lacking solutions, such as "express $\sqrt{2}$ as the ratio of two integers," or "find two real numbers x and y such that $x + y = 1$ and $xy = 1$ simultaneously," to the unsolvable problems implied by Gödel's Theorem, the potential solvability can be an issue that lurks in the background. We are faced with the paradoxical situation that the solution to a problem may be that there is no solution.

What kind of an answer will you accept as a solution? It is important to have in mind the purpose to which a presumptive solution will be put. (See [Uspenskii 1974], pp. 5–8, and [Wilf 1982].)

A so-called solution may be useless in certain situations and hence, not a solution at all.

[4] At the research level, "plugging back in" can have its own problems. See [Gautschi 1983].

Example: The expression of the determinant of an nxn matrix in terms of $n!$ monomials formed from the matrix elements is pretty useless in the world of scientific computation. One looks around for other ways and finds them.

Example: Most finite algorithm problems have a solution that involves enumerating all the possibilities and checking, but this brute force strategy is seldom a satisfactory solution and is certainly not an aesthetic solution.

Example: A differential equation may be solved by exhibiting its solution as an integral. But to a college undergraduate who has met up with integrals only in a previous semester, an integral is itself a problem and not a solution. An approximation to the solution of a differential equation may be exhibited as a table, a graph, a computer program or may be built into a chip. Is such a solution good enough in a particular situation?

Example: If the problem is to "identify" the sequence $1, 2, 9, 15, 16, \ldots$ will you accept a "closed" formula (query: what exactly do you consider as a closed formula?), a recurrence relation, an asymptotic formula, a generating function? A semi-verbal description? Do you want statistical averages or other properties? Will you try to find the sequence in *The Online Encyclopedia of Integer Sequences*? Or will you simply say that a finite sequence of numbers can be extended to an infinite sequence in an unlimited number of ways and chuck the problem out the window as ill-formulated? How would you even elaborate explicitly the verb "identify" so as not to chuck the problem?

Though a problem has been solved in one particular way, the manner of solution may suggest that it would be very nice to have an alternate solution. An interesting instance of this is the prime number theorem. Originally proved via complex variable methods, Norbert Wiener (and others) asked for a real variable proof. Since the statement of the prime number theorem involves only real numbers, the demand for such a proof was possibly a matter of mathematical aesthetics. A real variable proof was given by Paul Erdös and Atle Selberg in 1949, partly independently.

Is such and such really a solution? There are constructive solutions but, as already observed, a solution may be "constructive" in principle but in practice the construction would take too long to be of any actual use. (The dimensional effect or the n! effect.)

Then there are existential solutions in which the generic statement is "There exists a number, a function, a structure, a whatever, such that. . . . " The mathematician Paul Gordan (1837–1912), when confronted with Hilbert's existential (i.e., non-constructive) proof of the existence of a finite rational integral basis for binary invariants, asked "Is this mathematics or theology?" ([Reid 1970], pp. 34–37)

Example: The Mean Value Theorem asserts that given a function $f(x)$, continuous on (a, b) and differentiable on (a, b), there exists a ξ in (a, b) such that $f(b) - f(a) = f'(\xi)(b - a)$. Some students find this statement hard to take when they first meet up with it. The ξ appears mysterious.

Example: The famous Pigeonhole Principle: Given m boxes and n objects in the boxes where $n$is larger than m. Then there exists at least one box that contains more than one object. Who can deny this? This may lead to an existential solution. On this basis, for example, together with some tonsorial data, one can conclude that there are two people in Manhattan that have the same number of hairs on their head. Now find them. We have been assured that they can surely be located in a platonic universe of mortals.

Example: There exist irrational numbers x and y such that x^y is rational. Proof: Set $r = \sqrt{2}^{\sqrt{2}}$.Now if r is rational, then since $\sqrt{2}$ is irrational, the selection $x = y = \sqrt{2}$ works. On the other hand, if r is irrational, then set $x = r$ and $y = \sqrt{2}$. Since $x^y = (\sqrt{2}^{\sqrt{2}})^{\sqrt{2}} = (\sqrt{2})^2 = 2$, this selection works. One may ask: is this really a solution if we can't, with our present knowledge, decide whether r is or is not rational?

There are "probabilistic solutions" as, for example, the Rabin-Miller probabilistic test for the primality of a large integer. [Rabin 1980]

Then there are the "weak solutions." In 1934, Jean Leray proved that there is a weak solution to the incompressible Navier-Stokes equations. Is there only one such? The question appears to be still open. But what, in a few sentences, is a weak solution? If there is ambiguity about the very notion of a "solution," this is equally the case for a "weak solution." Technically, if L is a differential operator, and if u = f satisfies the equation Lu = g, then f is the solution. If so, then for all "test functions" ø,

(Lf, ø) = (g, ø), (,) designating an inner product. But if only the latter is true, f is said to be a "weak solution."

Since some problems are very difficult, or even unreachable with current mathematical theory and techniques, the notion of a weak problem, possessing weak solutions, has been introduced as a framework that allows existing mathematical tools to solve them. A strong solution is a weak one but often a weak solution is not a strong one, and the relation between the two notions is still the subject of intense research.

In a numerical problem, is a weak solution really a solution if it is not computable? Despite this limitation, the knowledge that a weak solution exists can have a have considerable impact.

Apparently, the meaning of the word "solution" can be stretched quite a bit. The elastic quality of mathematical terms or definitions is remarkable, and is often achieved through context enlargement.

There are cases where a problem has been turned into its opposite. Thus, the search for the dependence of Euclid's Fifth Axiom (the parallel axiom) on the other axioms, resulted in the unanticipated knowledge of its independence. The Axiom of Choice was hopefully derivable from the other axioms of set theory. It is now known to be independent of them. An instruction to prove, disprove, or prove that neither proof nor disproof is possible, is a legitimate, though a psychologically unpleasant formulation of a problem.

There are cases where a problem was felt to be solved, and then later was felt to be open, not because an error was found, but because there was a shift in the (unconscious) interpretation of what had been given. For this, read Imre Lakatos' classic discussio*n of the history of the Euler-Poincaré theorem. A very early version reads* $V - E + F = 2$ where V, E, and F are respectively the number of vertices, edges, and faces of a polyhedron. But just what kind of a 3-dimensional object is a polyhedron and what are its vertices, edges and faces? Lakatos' discussion chronicles the ensuing tug-of war—almost comic—between hypotheses and conclusions and the negotiations necessary so as to maintain a semblance of the original conclusion. This is known in philosophy as "saving the phenomenon." [Lakatos 1976].

6 When is a Proof Complete?

If the problem is to find a proof (or a disproof) of a conjecture, how does one know that that a purported proof is correct? Gallons and gallons of ink have been expended on this question as formulated generally. Are proofs stable over time? A half century after D'Alembert gave a proof of the Fundamental Theorem of Algebra, Gauss criticized it. A century after Gauss' first proof (he gave four), Alexander Ostrowski criticized it.

Is a proof legitimate if it is hundreds of pages long and would tire most of its human checkers? Is a proof by computer considered legitimate? The publicized proof by Thomas Hales of the Kepler sphere packing conjecture is said to require 250 pages of text and 3 gigabytes of programs. The mathematical community is itself split over the philosophical implications of the answers given to these and a myriad of similar questions. [Hales www]

For one criterion as to when a solution is a solution, when a proof is a proof, let's go, as bank robber Willie Sutton said he went, to where the money is. A recent answer to this question was formulated by the Clay Mathematics Institute which offers prizes of a million dollars for the solution of each of seven famous problems. The Clay criteria for determining whether a problem is solved are as follows.

(1) The solution must be published in a refereed journal.
(2) A wait of two years must ensue after which time if the solution is still "generally acceptable" to the mathematical community,
(3) the Clay Institute will appoint its own committee to verify the solution.

In short, a solution is accepted as such if a group of qualified experts in the field agree that it's a solution. This comes close to an assertion of the socially constructive nature of mathematics. The remarkable thing is the social phenomenon of (almost) universal, but not necessarily rapid, agreement , which has been cited as strengthening mathematical platonism. (See [Davis 1990], [Ernest 1998], [Rosental 2003].)

7 Applied Mathematics

In applied mathematics—and I include here both physical and social models—other answers to the basic question of this article can be put forward. Proofs may not be of importance. The formulation of adequate mathematical models and adequate computer algorithms may be all important. What may be sought is not a solution but a "good enough solution."

In introductions to applied mathematical and in philosophical texts, loops are often displayed to outline and conceptualize the process. The loops indicate a flow from

(a) the real world problem to
(b) the formulation of a mathematical model, to
(c) the theoretical consequences of the model, to
(d) the computer algorithm or code, to
(e) the computer output to
(f) the comparison between output and experiment.

Then back to any one of (b)–(f) at any stage. And even back to (a), for in the intervening time, the real world problem may have changed, may have been reconceived, or even abandoned.

In looking over these steps, it occurred to me that one additional step is missing from this standardized list. It is that (f) can lead to

(g) an action taken in the real world and to the responses of the real world to this action.

This omission might be explained as follows: at every stage of the process one must certainly simplify—but not too much, else verisimilitude will be lost. The responses of the real world are both of a physical and of a human nature, and the latter is notoriously difficult to handle via mathematical modeling. Hence there is a temptation to "put a diagrammatic wall" around (b) to (e) that emphasizes the mathematical portion as though mathematics gets done in a sanitized world of idealized concepts that does not relate to humans. Step (g) is often conflated with (f) and let go at that. Since we are living in a thoroughly mathematized world with additional mathematizations inserted by fiat every day that impact our lives in myriads of ways, it is vital to distinguish (g) and to emphasize it as a separate stage of the process.

What cannot be known in advance is how often these loops must be traversed before one says the problem has been adequately solved. Common sense, experience, the support of the larger community in terms of encouragement and funding may all be involved arriving at a judgment. And yet, one may still wonder whether steps (a)–(g) provide a sufficiently accurate description of the methodology of applied mathematics.

8 Some Historical Perspectives

One can throw historical light on the question of when a problem is solved. There are several ways of writing the history of mathematics. I'll call them the horizontal and the vertical ways. In horizontal history, one tries to tell all that was going on in, say, the period 400–300 B.C. or between 1801 and 1855. In vertical history, one selects a specific theme or mathematical seed, and shows how, from our contemporary perspective, it has blossomed over time. (See [Grattan-Guinness 2004].)

As a piece of vertical mini-history, consider the quadratic algebraic equations first met in high school. Such equations were "solved" by the Babylonians 4,000 years ago. But over the years, immense new problems came out of this equation in a variety of ways: higher order algebraic equations, the real number system as we now know it, complex numbers and algebraic geometries; group and field theory, modern number theory, numerical analysis.

Solving a polynomial algebraic equation of degree n once meant finding a positive rational solution. Today it means finding all solutions, real or complex together with their multiplicities and finding it either in closed form (rare) or by means of a convergent algorithm whose rate of convergence can be specified. But the generalizations of quadratic equations go further. Formal equations can be interpreted as a matrix or even as an operator equation in various abstract spaces. The equation $x^2 = 0$ trivially has only $x = 0$ as its solution when x is either real or complex. But this is not the case if x is interpreted as an $n$by n matrix: the nilpotent matrices solve this equation. And if you have the temerity to ask for all nilpotent operators in abstract spaces, you have raised a question without a foreseeable end.

A more recent example, of which there are multitudes. In 1959, Gelfand asked for the index of systems of linear elliptic differential equations on compact manifolds without boundary. The problem was solved in 1963 by Atiyah and Singer, and this opened up new ramifications with surprising features including Alain Connes' work on non-commutative geometry.

In the historical context, mathematical problems are never solved. Material, well established, is gone over and over again. New proofs, often simplified, are produced; contexts are varied, enlarged, united, and generalized. Remarkable connections are found. Repetition, reexamination are parts of the practice of mathematics.

9 A Dialogue on When is a Theory Complete

The original question as to when is a problem solved may be moved up a level to ask: when is a theory complete? Stephen Maurer, one of the MathPath faculty, provided me with a web discussion of this question he'd had with one of his most philosophical students. I present it here as Maurer sent it to me.

> Andy Drucker:
>
> "This question has been haunting me, and I know I shouldn't expect definite answers. But how do mathematicians know when a theory is more or less done? Is it when they've reached a systematic classification theorem or a computational method for the objects they were looking for? Do they typically begin with ambitions as to the capabilities they'd like to achieve? I suppose there's nuanced interaction here, for instance, in seeking theoretical comprehension of vector spaces we find that these spaces can be characterized by possibly finite 'basis' sets. Does this lead us to want to construct algorithmically these new ensembles whose existence we weren't aware of to begin with? Or, pessimistically, do the results just start petering out, either because the 'interesting' ones are exhausted or because as we push out into theorem-space it becomes too wild and wooly to reward our efforts? Are there more compelling things to discover about vector spaces in general, or do we need to start scrutinizing specific vector spaces for neat quirks—or introduce additional structure into our axioms (or definitions): dot products, angles, magnitudes, etc.?
>
> Also, how strong or detailed is the typical mathematician's sense of the openness or settledness of the various theories? And is there an alternative hypothesis I'm missing?"
>
> Stephen Maurer:
>
> "This is an absolutely wonderful question—how do mathematicians know when a theory is done—and you are right that there is no definitive answer. The two answers you gave are both correct, and I can think of a third. Your two answers were 1) we know it's done when the questions people set out to answer have been answered, and 2) we know it's done when new results dry up. My third answer is 3) we don't know when it's done.
>
> An individual probably feels done with a theory when the questions that led him/her to the subject are answered (answered in a way that he feels gives a real understanding) and he either sees no further interesting follow-up questions or can't make progress on the ones he sees. Mathematicians as a group probably feel it's done when progress peters out—the subject is no longer hot and it is easier to make a reputation in some other field that is opening up. (You called this attitude pessimistic, and I'm not so keen about it either, but it shows that math, like other subjects, is influenced by more than pure thought, and it means that mathematicians are trying to optimize results/effort.)

> But finally, history shows that fields are rarely ever done. Much later a new way of looking at an old field may arise, and then it's a new ball game. Geometry is an example. The study of n-dimensions was around long before vectors and dot products (there are books of n-dimensional theorems proved by classical Euclidean methods) but the creation of these vector ideas in physics led to a new blossoming of geometry.
>
> Another example is the field of matroids, in which I got my Ph.D. Matroids have been described as "linear algebra without the algebra." Concepts such as basis and independence make sense (and have the same theorems you have seen, such as that all bases have the same size) but there is no plus or scalar multiplication! Matroids were invented in the 1930s, for a different purpose than generalizing linear algebra, and lay fallow for some time. Then, starting in the 1960s, their general value was appreciated and they sprung to life for perhaps 30 years. We might have said that we thought linear algebra was done, but since matroids are a form of linear algebra generalization, we discovered it was not done.
>
> Now matroids are fairly quiet again; there are still papers published in the field, but the natural questions that occurred to people when the subject was fresh have been answered or people have mostly stopped trying. It has become, like linear algebra itself, a background theory that people apply when appropriate."

Examples of revitalization abound. At the end of the 19th century, it was thought that invariant theory was finished and that Hilbert's work had killed it off. But it lives on. Where is nomography today? Its theoretical heyday seems to have been in the work of Maurice d'Ocagne [d'Ocagne 1899], but it lives on in engineering circles. See also [P. Davis 1995] for another example of revitalization in geometry.

Reading the Drucker-Maurer dialogue recalled to my mind that Felix Klein (1849–1925) and John von Neumann (1903–1957) emphasized other sources of revitalization. Felix Klein:

> "It should always be required that a mathematical subject not be considered exhausted until it has become intuitively evident...." ([Kline 1972], p. 904)

By Klein's criterion, and considering contemporary proofs that require hundreds of pages or are done with a computer assist, it would appear that many mathematical subjects have a long life ahead of them before they become intuitively evident.

Von Neumann's answer contains a cautionary message which I, as an applied mathematician, appreciate. I reproduce a short portion of his article.

> "As a mathematical discipline travels far from its empirical source, or still more, if it is a second and third generation only indirectly inspired from ideas coming from 'reality,' it is beset with very grave dangers. It becomes more and more purely aestheticizing, more and more purely *l'art pour l'art*. This need not be bad, if the field is surrounded by correlated subjects, which still have closer empirical connections, or if the discipline is under the influence of men with an exceptionally well-developed taste.
>
> But there is a grave danger that the subject will develop along the line of least resistance, that the stream, so far from its source, will separate into a multitude of insignificant branches, and that the discipline will become a disorganized mass of details and complexities.

> In other words, at a great distance from its empirical source, or after much 'abstract' inbreeding, a mathematical subject is in danger of degeneration. At the inception the style is usually classical; when it shows signs of becoming baroque the danger signal is up. It would be easy to give examples, to trace specific evolutions into the baroque and the very high baroque, but this would be too technical.
>
> In any event, whenever this stage is reached, the only remedy seems to me to be the rejuvenating return to the source: the reinjection of more or less directly empirical ideas. I am convinced that this is a necessary condition to conserve the freshness and the vitality of the subject, and that this will remain so in the future." [von Neumann 1947]

10 A Possible Example of Renewal from the Outside

It may be invidious to mention a specific example of exhaustion of a field when there are people working very happily in it. But the following example and opinion is in the open literature. (See [Mumford 2000].) Classical mathematical logic, which proceeds from Aristotle through Frege, Russell & Whitehead, Tarski, and later, has lost its connection to reality and has produced mathematical monsters. The change that is suggested is to develop logics that build in theories of probability. There currently exist a number of probabilistic logics, but they are not entirely successful. Some have even said: construct logics that build in "intent" in the sense of the mathematical philosophy of Edmund Husserl.

11 Implications for Mathematical Education

What are some of the pedagogic implications of the discussions of this article?

Normally, the average student thinks of a mathematical problem as something where one arrives at a single answer as quickly as possible and then moves on to the next assigned problem. Brighter students—those who will go further with mathematics—should be encouraged to think of a problem as never really finished.

Other ways of looking at the problem may emerge and yield new insights. It is also important to examine a problem in relation to other parts of mathematics as well as to the historical and cultural flow of ideas in which it is embedded.

Discovering a sense in which a solved problem is still not completely solved but leads to new and profound challenges, is one important direction that mathematical research takes. To be fully alive in the world of mathematics is to be constantly aware of this possibility.

Finally, alluding to my MathPath experience that gave rise to this article, taking a student's question seriously can be fruitful for both the student and the professor. "Out of the mouths of babes and sucklings have I found strength."

Acknowledgements I wish to acknowledge my debt to George Rubin Thomas, the director of MathPath 2004, to its faculty and, in particular, to Elizabeth Roberts who asked the question. Thanks also to Stephen Maurer of Swarthmore College and the MathPath faculty who provided me with the dialogue that ensued when his graduate student Andy Drucker asked a higher-level question.

Thanks to the following mathematical friends who have also found the question stimulating: Bernhelm Booss-Bavnbek, Chandler Davis, Ernest S. Davis, Reuben Hersh, Yvon Maday, David Mumford, Kati Munkacsy, Kay L. O'Halloran. I have built their responses into this article. And finally, thanks to Bonnie Gold and Roger Simons for providing me with a number of textual and editorial suggestions and for including this article in their book.

The day hardly passes in which I do not receive further reponses and ramifications from additional friends. I assert firmly that I will never know when this article will really be finished. "The Song Is Ended but the Melody Lingers On"—Irving Berlin.

Bibliography

[Clay] Clay Institute criteria: http://www.claymath.org/millennium/Rules_etc/

[Davis 1990] Chandler Davis, "Criticisms of the Usual Rationale for Validity in Mathematics," pp. 343–356 in *Physicalism in Mathematics*, A.D. Irvine, ed., Kluwer, 1990.

[Davis 1987] Philip J. Davis, "When Mathematics says No," in *No Way*, Philip J. Davis and David Park, eds., W.H. Freeman, 1987.

[Davis 1995] ——, "The Rise, Fall, and Possible Transfiguration of Triangle Geometry: A Mini-history," *American Mathematical Monthly* 102 (1995), pp. 20–214.

[Ehrich 2001] Sven Ehrich, "Stopping Functionals for Gaussian Quadrature Formulas," *J. Comp. and Appl. Math.* 127 (2001), pp. 153–171.

[Ernest 1998] Paul Ernest, *Social Constructivism as a Philosophy of Mathematics,* SUNY Albany, 1998.

[Finkelstein/Levin 2004] Michael Finkelstein and Bruce Levin, "Stopping rules in clinical trials," *Chance* 17 (2004), pp. 39–42.

[Gautschi 1983] Walter Gautschi, "How and How Not to Check Gaussian Quadrature Formulae," *BIT* 23 (1983) pp. 209–216.

[Grattan-Guinness 2004] Ivor Grattan-Guinness, "History or Heritage? An Important Distinction in mathematics and in mathematics education," *Amer. Math. Monthly* 111 (2004), pp. 1–12.

[Hales www] Thomas Hales' proof: http://www.maa.org/devlin/devlin_9_98.html.

[Mumford 2000] David Mumford, "The Dawning of the Age of Stochasticity," *Rend. Mat. Acc. Lincei* 9 (2000), pp. 107–125.

[Kline 1972] Morris Kline, *Mathematical Thought from Ancient to Modern Times*, New York, Oxford University Press, 1972.

[Lakatos 1976] Imre Lakatos, *Proofs and Refutations*, Cambridge Univ. Press, 1976.

[Lakoff/Núñez 2000] George Lakoff and Rafael Núñez, *Where Mathematics Comes From: How the Embodied Mind Brings Mathematics into Being,* Basic Books, 2000.

[von Neumann 1947] John von Neumann, "The Mathematician," in *The Works of the Mind,* Robert B. Heywood, ed., Univ. Chicago Press, 1947. Reprinted in *Musings of the Masters*, Raymong G. Ayoub, ed., Mathematical Association of America, 2004, pp. 169–184.

[d'Ocagne 1899] Maurice d'Ocagne, *Traite de Nomographie,* G. Villars, 1899.

[O'Halloran 2005] Kay L.O'Halloran, e-mail correspondence. Also: *Mathematical Discourse: Language, Symbolism, and Visual Images*. Continuum, London and New York, 2005.

[Rabin 1980] M. O. Rabin, "Probabilistic Algorithm for Testing Primality," *J. Number Th.* 12 (1980), pp. 128–138.

[Reid 1970] Constance Reid, *Hilbert*, Springer Verlag, 1970.

[Rosental 2003] Claude Rosental,"Certifying Knowledge: The Sociology of a Logical Theorem in Artificial Intelligence," *American Sociological Review*, 68 (2003), pp. 623–644.

[Uspenskii 1974] V.A. Uspenskii, *Pascal's Triangle*, University of Chicago Press, 1974.

[Wilf 1982] Herbert Wilf, "What is an answer?" *Amer. Math. Monthly*, 89 (1982), pp. 289–292.