

The relevance of the irrelevant beginning

Philip J. Davis*

Division of Applied Mathematics, Brown University, Providence, RI 02912, USA

*Corresponding author's e-mail address: philip_davis@brown.edu

Published online: April 28, 2014 (version 1)

Cite as: Davis P. J., ScienceOpen Research 2014 (DOI: 10.14293/A2199-1006.01.SOR-MATH.6G464.v1)

Reviewing status: Please note that this article is under continuous review. For the current reviewing status and the latest referee's comments please click [here](#) or scan the QR code at the end of this article.

Primary discipline: History & Philosophy

Associated discipline: Applied mathematics, General mathematics

Keywords: Sequences, Convergence, Computation, Induction, Mathematical philosophy

ABSTRACT

There are many paradoxes associated with the concept of infinity. I would like to set forth and comment on one which, though well-known in principle, I have not seen stated in the following form.

A PARADOX

Given a sequence of real numbers, x_1, x_2, x_3, \dots , Determine whether the sequence is convergent or divergent. If it is convergent, what is its limit value? For any finite integer, n , the determination of convergence cannot be made theoretically on the basis only of one's knowledge of $x_1, x_2, x_3, \dots, x_n$. Since n can be selected arbitrarily large, it follows that the question cannot be answered at all. Does this imply that the concept of convergence is meaningless?

On the other hand, every day of the week, practical scientific computations employing iterative methods productively accept a "convergent value" on the basis of a finite number of values. (Or the investigator may decide that the iterative method employed is divergent.)

I call this the "Paradox of the Irrelevant Beginning" because while the knowledge of the first n terms of the sequence appears to be theoretically irrelevant as regards answering the question, yet "answers" to the question are accepted. What, knowledge, then, is relevant? Here is a conflict between theoretical mathematics and applied mathematical practice. They appear to be at loggerheads. Is the paradox above mere verbal confusion or is there something in it that is deeper?

APPLIED MATHEMATICS

At the very elementary computational level, iterative methods are employed; for example, fractions, such as $1/3$ and $4/19$, are converted to (mostly) approximate binary or decimal forms and then are manipulated (mostly) in those

approximate forms.¹ These iterations and those for elementary functions, such as \sqrt{x} , and iterations, such as $a_{n+1} = 1/2(a_n + x/a_n)$, are well-understood. When it comes to the iterations employed in, say, the numerical solution of nonlinear partial differential equations, this may not be the case.

A leading investigator, AZ, who has much experience with the numerical solution of the Navier–Stokes equations of fluid dynamics, responded as follows to several questions I put to him:

PJD: Do I assume correctly that in writing computer algorithms for your professional work, you have used iterative methods?

AZ: Yes a lot. This is quite mandatory to solve nonlinear problems that are the real ones coming from the applications. In this case the iterations are based on the Newton algorithm of Picard fixed point.... These iterations often require the solution of a linear part. For the resolution of the linear part we often use also iterative methods, these are based on conjugate gradient approaches or Krylov methods. For the resolution of linear problem through iterative methods, preconditioners are often involved. These are based on simplified versions of the original problem, they are easy to solve and allow to diminish the number of iterations

PJD: If this is the case, what criteria do you employ for terminating the iterations and outputting a result?

AZ: The most simple is the size of the increment ... but this is (of course) too simple for real problems. We prefer to look at the size of the residual, the most efficient method are based on good norms which are not so easy to derive ... but this is the price to pay

PJD: In the computations involving iterative methods do you have any *a priori* theoretical information as to the rate of convergence of the algorithms?

¹ Thus, under "normal usage" MATLAB yields, e.g., $1/2 + 1/3 - 5/6 = -10^{-16}$.

AZ: Yes, we generally have something. The more we have the most efficient is the stopping criteria based on the size of the increment. In most real problem, we do not use it much since it is too rough

PJD: Do you employ “speed up” methods to improve the accuracy? If so, what knowledge do they depend upon?

AZ: As I said before we use a lot of preconditioners, they are more based on what we know about real life problems and what we expect ... it is quite know how related ... when you have experienced large number of iterations on one problem, you try something to diminish the number of iterations. The knowledge of 3 is thus rarely used.

A few comments on AZ’s response. After how many iterations does the investigator tell the computer to halt and exit with a value? There is a large numerical analysis literature on this question that provides rules of thumb. I am aware of at least 13 different “stopping,” “halting,” or “termination” criteria. One of the very simplest is: iterate until $x_{n+1} - x_n$ is about the size of the machine 0, having first set an upper limit to the number of allowable iterations. I am unaware of a taxonomic study of termination criteria [1]. If we have an idea of the rate of convergence, linear, quadratic, etc., then “speedup” or “extrapolation to infinity” methods are sometimes useful. And, of course, the critical problem of stable versus unstable iterations lurks in the background. As regards “rules of thumb,” it is part of the job of applied mathematics to extract more theoretical information out of the equations so that more efficient methods can be discovered.

There are various problems with the strategy of “eyeballing” the output from $n = 1$ to $n = 10$ or from $n = 1$ to 10^{10} or subjecting it to various statistical tests or termination criteria. What happens when dealing with very slowly convergent sequences? Ralph Boas Jr. provided an amusing example when he asked how many terms are necessary to compute two correct decimal places of the convergent sum:

$$S = \sum_{n=2}^{\infty} (1/(n \log(n \log \log(n))^2))$$

The number of terms that need to be added is 10^{87} which cannot be done in the lifetime of the universe [2]. (Of course, this does not rule out the possibility that an equivalent representation for S might be found leading rapidly to a correct two-figure answer.) Boas’ example may be an amusing but an exaggerated case; on the other hand, there are numerous very slowly convergent series arising in mathematical physics and devices for summing them efficiently. Another possibility is that one is dealing with an asymptotic (divergent) series where “convergence” takes place as long as one does not take too many terms [3].

The applied mathematician who computes, for example, certain aspects of fluid dynamics by a discretized scheme that invokes iteration, does indeed use the information yielded by the initial terms of a hopefully convergent sequence. What

may then come into play is collateral information (perhaps from aerodynamical experiments) or other pieces of evidence that point to the credibility of his accepted (i.e., shut-off) value. His algorithm, together with its shut-off criterion, constitutes a model of a natural phenomenon and the model may stand or fall on the basis of nonmathematical criteria.

Thus, the computational strategy of applied mathematics involving iterations is based on the supposition (or the faith) that the particular iterative method employed will converge with sufficient accuracy within an amount of human time that is agreeable to the investigator and be consistent with the phenomenon as experienced.

We may consider the initial portion of an iterative sequence as simply a finite set of numbers. Applied mathematics constantly formulates conjectures and bases conclusions on finite sets of data – often experimental. The theory of data analysis involves curve-fitting, statistical criteria, etc. The theory goes back to Gauss, if not earlier, and is now embodied in a large variety of computer packages and chips.

PURE MATHEMATICS

Yet the paradox of the irrelevant beginning goes beyond applied mathematics and into the methodology of pure mathematics. If by “methodology” we admit the formulation of conjectures, then consider the following statement regarding the “normality”² of the digits of pi (π):

Pi certainly seems to behave this way. In the first six billion decimal places of pi, each of the digits from 0 through 9 shows up about six hundred million times. Yet such results, conceivably accidental, do not prove normality even in base 10, much less normality in other number bases. In fact, not a single naturally occurring math constant has been proved normal in even one number base, to the chagrin of mathematicians. While many constants are believed to be normal – including pi, the square root of 2, and the natural logarithm of 2 – there are no proofs [4]

The evidence of six billion decimal places certainly strengthens the conjecture – if only psychologically – that π is a normal number. We can locate this research and its accompanying statement generically within the George Pólya heuristic of discovery: if A implies B and B is true, then A is more likely to be true [5].³

Concepts simpler than normality form the bases of conjectures based on finite numerical evidence. For example, it is not known at the moment of writing whether the

² Normality of a real number α expressed in base b with digits d_0, \dots, d_{b-1} means that all these digits occur in α and are uniformly distributed.

³ In probabilistic terms: if $\text{prob}(B|A) > \text{prob}(\text{not } B|A)$ and $\text{prob}(A) > 0$, then $\text{prob}(A|B) > \text{prob}(A)$.

Euler–Mascheroni constant γ (gamma) defined by $\gamma = \lim_{x \rightarrow \infty} \left(\sum_{k=1}^x \frac{1}{k} - \ln x \right)$ is rational or irrational. What is known, but not on the examination of the digits of $\gamma \approx 0.57721\ 56649\ 01532\ 86060\ 65120\ 90082\ 40243\ 10421\ 59335\dots$ but based on alternate representations of γ , is that if $\gamma = a/b$, where a and b are integers reduced to lowest terms, then b must exceed $10^{200,000}$. Again, this strengthens the conjecture that γ is irrational.

Perhaps the most famous example of finite computations leading to a strengthened conjecture relates to the Riemann hypothesis, where it has been shown that the first 2.4×10^{13} nontrivial zeros of the ζ function lie on the critical line. Of course, the Devil’s Advocate can point out that there are numerous examples in which some conjectured condition is “yes” for the first gazillion cases and then becomes “no” hereafter.

There is another kind of reliance on initial segments that occur within pure mathematics. Frequent mathematical questions on intelligence tests are of the following sort: what is the next number in the sequence 1, 2, 4, 8, ...? Strictly speaking, any number at all would do for which a compatible polynomial formula is easily found. But what answer is expected by the graders of the intelligence test? On the other hand, those concerned with mathematical heuristics have found that it is useful and makes good sense to create the *Online Encyclopedia of Integer Sequences* that will attempt to identify a finite sequence from among a large and growing database of sequences that have arisen in theoretical work.

Examples:

- (1) 1, 2, 4, 8, ... The *Encyclopedia* listed 920 answers. Under the reference A088274, we find “numbers n such that $10^n + 7$ is prime.” The next two numbers in the sequence are 9 and 24.
- (2) I entered the sequence 1, 6, 7, 18, plucked “at random” from my brain and I was lucky. The *Encyclopedia* responded with only “Representation degeneracies for Neveu–Schwarz rings,” and carried the sequence forward as 29, 59, 92, ...
Thus, in these situations answers were provided on the basis of initial data.
- (3) On the other hand, I entered 1, 2, 4, 15, 16 and drew a blank.

CONSTRUCTIVISM VERSUS EXISTENTIALISM

Let us look at the matter from the point of view of a new student studying elementary calculus. Early on, the student learns that the sequence $x_n = r^n$ converges as $n \rightarrow \infty$ if and only if $-1 < r < 1$ or $r = 1$. The student learns that the limit as $h \rightarrow 0$, of $(\sin h)/h = 1$, etc. The student will surely be exposed to proofs of these results. In elementary calculus, we pose exercises for the student such as to find $\lim_{n \rightarrow \infty} n \cdot \left(\frac{an^2 + bn + c}{fn^2 + gn + k} \right)$ and such exercises can be done on a small piece of paper. They certainly do not require assistance from a digital computer.

An intuitive notion of convergence as “getting closer and closer” to some number predominates, backed up possibly by a graphical figure, and the easily believed rules of elementary combination such as the limit of a sum is the sum of the limits, etc.

Already Archimedes (c. 200 BCE) in his famous determination of the value of π , using inscribed and circumscribed regular polygons, came out with inequalities that enclosed π between two rational numbers: $3\ 10/71 < \pi < 3\ 1/7$. This result set an ideal toward which all subsequent iterators aspire: sharp error bounds. As the subject developed, mathematics required more sophistication and precision than mere intuitive notions. Still, there was no early attempt to produce a formal definition of what convergence was all about. Such a definition, which appears late in the history of mathematics (A.L. Cauchy, *Cours*, 1821), is now absolutely basic to analysis and it is famous for its implied precision and infamous for a certain vagueness. A contemporary form is:

The sequence $\{a_n\}$ has the limit a if given any $\varepsilon > 0$, we can find or (or there exists) an N such that $|a_n - a| < \varepsilon$ whenever $n \geq n_0$.

Still more definitive in a certain sense is the notion of a “Cauchy sequence” $\{a_n\}$:

Given any ε , we can find (or there exists) an n_0 such that $|a_m - a_n| < \varepsilon$ whenever m and $n \geq n_0$.

Notice the generality that is implied here. Whereas in student exercises, a closed form expression is set forth to work on, it is not the case here. Note also that the second criterion does not even specify what the limiting value is. By an act of our imagination, we create in our minds an infinite sequence of numbers, not specifying in any additional sense what the numbers are, but to which we say we can apply these famous ε , n_0 definitions of convergence.

The vagueness in the Cauchy-type definitions is often troublesome for students. What is meant by the expressions “there exists” or “we can find?” How can we find? And again, there are the pesky words “given” or “let.” What is “any number ε ” and how does one “give an ε ?” Does the word “any” in the expression “given any ε ” imply an unlimited number of acts of “giving”? Do such acts take place in real time? Not only students but also philosophers of mathematics have found these questions troubling.

Still, on the basis of the definitions of convergence previously mentioned, one goes on to prove algebraic permanences of convergence (e.g., the sum of two convergent sequences is convergent) but more complicated statements, such as if $\{a_n\}$ converges to the value a , then the arithmetic means of $\{a_1, a_2, \dots, a_n\}$ converge to a . Mathematical analysis is based on such material and the corresponding material for infinite series. One may dub this kind of theorematic “mathematical existentialism” for it postulates that by a pure mental act the existence of infinite sequences without further specifications or descriptions can be “brought up” and operated upon.

Mathematical existentialism is normative in mathematical practice, but it has run into objections first posed by the Dutch mathematician and philosopher L.E.J. Brouwer. It has been worried about by Hermann Weyl, and elaborated into “constructivist mathematics” by Brouwer’s followers such as Errett Bishop [6]. Thus (without following any particular “established” philosophic opinion or details), we may designate two attitudes toward mathematical practice: the existential and the constructive. I use these terms designating the two categories loosely in a personal, Humpty Dumpty sense, admitting that there is overlap, considerable ambiguity, and vague boundaries between them. One difficulty may be cited: a concept may have several different representations lying in different categories, for example, the square root of 2 ($\sqrt{2}$), thought of as an infinitely long decimal 1.4142135623731, residing without further specification in the world of the decimal representations of all real numbers is existential, while the rapidly convergent iteration:

$$a_0 = 1; a_{n+1} = (1/2)(a_n + 2/a_n), \lim_{n \rightarrow \infty} a_n = \sqrt{2}$$

can be thought of as a constructive representation of $\sqrt{2}$. Within the theory of quadratic number fields such as $Q(\sqrt{2})$: $\{a + b\sqrt{2}\}$, a, b , rational, the symbol $\sqrt{2}$ together with the rule, $(\sqrt{2})^2 = 2$, need not even have any numerically infinite implications.

If a sequence is given by a closed mathematical formula or an algorithm in some computer language, we may be able to answer the question of convergence by a penetrating study of the structure of the formula or of the algorithm. Over the centuries, various criteria and strategies or “tricks” or have been devised to test for convergence, and these, combined with various rules of combination, have been built into computer packages that have symbolic potentialities. Thus, using such a package (e.g., MATHEMATICA), we may pose the question what is the limit as $n \rightarrow \infty$ of $(an^2 + bn + c)/(fn^2 + gn + k)$ and obtain an answer.

These considerations lead immediately to the question of what are allowable closed formulas? At the pragmatic level, we may answer that a closed formula is one that is currently built into computational packages such as MATLAB or MAPLE. This fits in nicely with the conception of a mathematical function in the days of Euler (c. 1750). This pragmatic answer does not fit in with the currently promulgated definition that a function is an association of every element of a set A with a unique element of a set B and says nothing about how we are to conceive of or create such an association [7].⁴ Thus, contemporary computational practice is Eulerian in spirit, and if in the future the computer dominates mathematical

developments, the spirit of Euler will rise above that of Cauchy.⁵

The notion of computability has now generated extensive theories of computable numbers and computable functions centering around the names of Gödel, Turing, Chaitin, et al. Such theories appear, *prima facie*, to be constructivist. But they soon wander off and create entities such as “inaccessible numbers” that for me exhibit a dubious ontology. Call this mathematical theology if you will. In any case, there is a considerable disconnect between what these theories provide and what applied mathematicians, in fact, have been able to utilize.

THE PARADOX CONSIDERED

One can say that the question of convergence of a specifically defined sequence belongs in the realm of constructive mathematics. On the other hand, the paradox of the irrelevant beginning, as phrased at the beginning of this article, is located within “existential mathematics.” We say “given an infinite sequence of numbers” or “consider a sequence of numbers” What do these directives mean? How do we “give,” “consider,” or “let” and operate with these “givens”?

Thus, the paradox, as phrased initially, has imported into existential mathematics a notion which originated within constructive mathematics and which, in real world applications, is still located there. In the mind and terminology of Aristotle, this is an error of “*metabasis*”: the transplantation of knowledge and methods from one area or category to another [8]. Despite Aristotle, *metabasis* is frequent and even desirable. While the employment of an discretized iterative computer algorithm to solve a set of equations is located within constructive mathematics, the truth of the matter is that in certain scientific computer codes, the algorithm may be so complex that it has defied a theoretical study as regards convergence (or indeed as regards of any of the typical questions raised by pure theoretical mathematics).

The computational practice of applied mathematics: iterate and exit with a value has, on the one hand, been forced onto the investigator by a lack of theoretical knowledge and on the other hand it has been made entirely plausible by many past numerical experiences showing that the practice can be successful and productive.

THE PROBLEM OF INDUCTION

A final word. The matter considered here is a good example of the practical and philosophical *problem of induction*: what can be concluded in a general way from limited experience. In the discussions given here, induction occurs at two levels. On the one hand, in each particular case we are inducing the ultimate behavior from the terms we have computed. On the

⁴ An “association,” as in category theory, appears as an undefined term.

⁵ Symbolic computer calculations as opposed to the numerical have increased in practice and prominence, e.g., in both computational algebra, geometry, and theoretical physics.

other hand, at the meta-level, we are inducing that certain termination conditions, etc. that have worked on previous problems will work on the one at hand.

Discussions of the general problem of induction are many: on the philosophical side ranging from David Hume to Karl Popper and to more recent philosophers without in any way dispelling the mystery [9]. On the applied mathematical side there is Bayesian analysis, etc.

In 1926, the brilliant Frank Ramsey (who, alas, died young) wrote “we are all convinced by inductive arguments, and our conviction is reasonable because the world is so constituted that inductive arguments lead on the whole to true opinions.” Why should this frequently be the case is worthy of speculation.

ACKNOWLEDGMENTS

This article was inspired in part by the recent book by Graham Oppy [10], which contains discussions of a large number of paradoxes of infinity. I wish also to thank Ernest S. Davis for his perceptive criticism and suggestions.

REFERENCES

- [1] Arioli M, Duff I, Ruiz D. Stopping criteria for iterative solvers. *SIAM J Matrix Anal Appl.* 1992;13(1):138–144. doi:10.1137/0613012
- [2] Boas RP. Partial sums of infinite series, and how they grow. *Am Math Mon.* 1977;84(4):237–258.
- [3] Caliceti E, Jentschura U, Meyer-Hermann M, Soff G, Weniger EJ. Slowly convergent and divergent series in theoretical physics; University of Bologna, Preprint; 2004.
- [4] Bailey DH, Crandall RE. On the random character of fundamental constant expansions, *Exp Math.* 2001;10(2):161–320.
- [5] Polya G. *How to solve it: a new aspect of mathematical method.* Princeton (NJ): Princeton University Press; 1945.
- [6] Bridges D, Richman F. *Varieties of constructive mathematics.* Cambridge (UK): Cambridge University Press; 1987.
- [7] Kleiner I. Evolution of the function concept. *College Math J.* 1989;20:282–300.
- [8] Funkenstein A. *Theology and the scientific imagination.* Princeton (NY): Princeton University Press; 1986.
- [9] Diettrich O. A constructivist approach to the problem of induction. *Evol Cogn.* 1995;1(2):1–25
- [10] Oppy G. *Philosophical perspectives on infinity.* New York (NY): Cambridge University Press; 2006.

COMPETING INTERESTS

The author declares no competing interests.

PUBLISHING NOTES

© 2014 Philip J. Davis. This work has been published open access under Creative Commons Attribution License **CC BY 4.0**, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. Conditions, terms of use and publishing policy can be found at www.scienceopen.com.

Please note that this article may not have been peer reviewed yet and is under continuous post-publication peer review. For the current reviewing status please click [here](#) or scan the QR code on the right.



scienceOPEN.com
research+publishing network