Cauchy's Famous Wrong Proof

The following example stems from my own teaching experience. Once, when I came to the topic of sequences and series of functions while teaching an undergraduate analysis class, I realized that the book had done a particularily poor job on this topic (to protect the guilty, no reference will be given). Thus something had to be done. Since I had been rereading Imre Lakatos's delightful little book, *Proofs and Refutations*, I decided to see if his analysis of Cauchy's famous wrong proof could be adapted to the classroom.[1]

In presenting the topic of sequences and series of functions, I began, as always, with a goodly supply of carefully chosen examples, drew pictures of some of them, and left others for homework. After noting how nicely the examples behaved, I coaxed the following observation out of the students.

THEOREM. A convergent series of continuous functions converges to a continuous function.

After congratulating my students for making this brilliant conjecture, I pulled Cauchy's *Cours d'Analyse* (1821) from my briefcase, attributed the theorem to him, and then presented his proof. This dusty tome lent authority to the argument to be given below, and the students were pleased that I had taken the trouble to go back to original sources. As often as appropriate, I take relevant books to class, and after explaining how they bear on the material, I pass them around for the students to look at. If that is impossible, I try to have excerpts on overhead transparencies to show. The students appreciate this.

Before giving the proof, some notation needs to be introduced. Cauchy is dealing with a series of functions whose sum is the function s. The n^{th} partial sum of this series is denoted s_n, and the remainder by r_n. Today all of this is usually presented in terms of sequences, but I wanted to follow Cauchy fairly closely. Consequently, I read Cauchy's proof to them and wrote it on the board in translation:

> When the terms of the series contain the same variable x, and this series is convergent, and its different terms are continuous functions of x, in the neighborhood of a particular value assigned to this variable; and s_n, r_n and s are again three functions of the variable x, of which the first is evidently continuous with respect to x in the neighborhood of the particular value in question. This assumed, let us consider the increases that these three functions receive when one increases x by an infinitely small quantity. The increase of s_n will be, for all possible values of n, an infinitely small quantity; and that of r_n will become insensible at the same time as r_n [sic], if one assigns to n a very considerable value. Hence, the increase of the function s can only be an infinitely small quantity. From this remark, one deduces immediately the following proposition [i.e., the Theorem above].[2]

[1] Imre Lakatos, *Proofs and Refutations*, edited by John Worrall and Elie Zahar, Cambridge University Press, 1976, pp. 123–141.

[2] Augustin Cauchy, *Cours d'Analyse*, 1821, p. 120. Reprinted in his *Œuvres*, II, 3.

This file, cauchy-wrng-pf.tex, was typeset using TEX on May 30, 1996 at 0:00 A.M.

Since the students were familiar with Weierstrassian ϵ-δ techniques, I took the time to carefully formulate Cauchy's proof in the modern language that they were learning to understand.

PROOF: Let $\epsilon > 0$ be given. Then
(1) Since each function is continuous, their partial sum, $s_n(x)$, is continuous,

$$\exists\delta\,\forall a \quad |a| < \delta \Rightarrow |s_n(x+a) - s_n(x)| < \epsilon.$$

(2) Since the series converges at x,

$$\exists N\,\forall n > N \quad |r_n(x)| < \epsilon.$$

(3) Since the series converges at $x + a$,

$$\exists N\,\forall n > N \quad |r_n(x+a)| < \epsilon.$$

Thus:

$$\begin{aligned} &|s(x+a) - s(x)| \\ &\quad = |s_n(x+a) + r_n(x+a) - s_n(x) - r_n(x)| \\ &\quad \leq |s_n(x+a) - s_n(x)| + |r_n(x)| + |r_n(x+a)| \\ &\quad \leq 3\,\epsilon \end{aligned}$$

Hence the function s is continuous. [It is pedantic to insist on ending with precisely "ϵ." Why make the mathematics even more mysterious for the student.] Q.E.D.

By careful planning, the class ended just as the proof did, and I was relieved that there were no questions after class. The next day the students were upset, for they had done their homework (this was a good class) and observed that some of the examples I had given (and left the graphs as exercises) contradicted Cauchy's Theorem. But they were ready to do mathematics.

I asked about the counterexample they had discovered in their homework:

$$\sin(x) - \frac{1}{2}\sin(2x) + \frac{1}{3}\sin(3x) - \frac{1}{4}\sin(4x) + \cdots$$

Were all the terms of the series continuous functions? Did the series converge? Was the limit function really discontinuous? "Yes, yes, yes." they said. Well then, what about the theorem? Cauchy published it in his *Cours d'Analyse*, so it must be correct, right? "Yes," they readily agreed. They had also accepted the proof when it was presented in class, for it seemed correct to them. They were puzzled. Something was wrong, but what?

I asked if they had examined the proof to see if anything was wrong with it. No, that had not occurred to them. So I suggested that we should look at the proof carefully.

Imre Lakatos makes the argument that it was in the mid-nineteenth century that mathematicians made the same advance as my students were now making: When a proof is wrong, do not just abandon it, but analyze it carefully to see if there are any "hidden hypotheses" that would make it correct. Lakatos took this phrase from a student of Dirichlet, Philipp Ludwig von Seidel (1821–1896), who used it in 1847 when he took the steps that my students were now ready for.

Thus, we shall now analyze Cauchy's proof. In step (1) above we need to realize that δ depends on ϵ, x, and n. To make this explicit, we shall write $\delta(\epsilon, x, n)$. Now, in step (2), N depends on ϵ and x, so we write $N(\epsilon, x)$. However, in step (3), N depends on ϵ, x, and ALSO on a. Using the same notation, we express this by $N(\epsilon, x+a)$. Now comes the critical observation. To make Cauchy's proof work, we need an integer M bigger than $N(\epsilon, x)$ and simultaneously bigger than $N(\epsilon, x+a)$ for each a whose absolute value is less than $\delta(\epsilon, x, n)$. Thus we must know that

$$M = \mathrm{Max}_t N(\epsilon, t)$$

exists for all ϵ, i.e., that M does not depend on x. Consequently, the additional hypothesis that we need is the following:

$$\forall\epsilon > 0\ \exists M\ \ \forall n > M\ \ \forall x \quad |r_n(x)| < \epsilon.$$

This is the definition of uniform convergence, and is precisely what is needed to make Cauchy's theorem correct and the proof work.

What I have done here is to motivate the definition of uniform convergence. Had I just written it down in the usual definition-theorem-proof style of modern mathematics, it would appear to be very much ad hoc. The historical presentation allows the student to see the true origin of the concept. As Lakatos has observed, the correct concept is generated by the incorrect proof. *This is one case where I feel that a historical presentation is absolutely necessary to the understanding of the material.*

You may object that this type of presentation takes too much time, for it did take two whole class periods. But that is not so. The time was well spent. Presenting the wrong proof and then analyzing it to see what additional hypotheses are needed takes far less time than presenting an ad hoc definition, trying (probably unsuccessfully) to explain it, and then finally giving the proof. With my presentation there is no need to give a correct proof after the definition has been discovered; that is an easy exercise for the student. In fact, my students said there was no reason for me to write out a new proof, for they had a deep understanding of how it works. Moreover, with this presentation the students have also learned more. The opportunity to analyze an incorrect proof builds both confidence and skepticism (students must learn that books may contain errors). More importantly, it shows them where theorems come from: We make conjectures, attempt proofs, analyze them, and refine them. It also shows the importance of definitions, showing that they are carefully chosen, not things arbitrarily written down just before a proof. I trust you will agree with my assessment that without giving this historical presentation, the student's understanding of the concept of uniform convergence would be severely hampered.

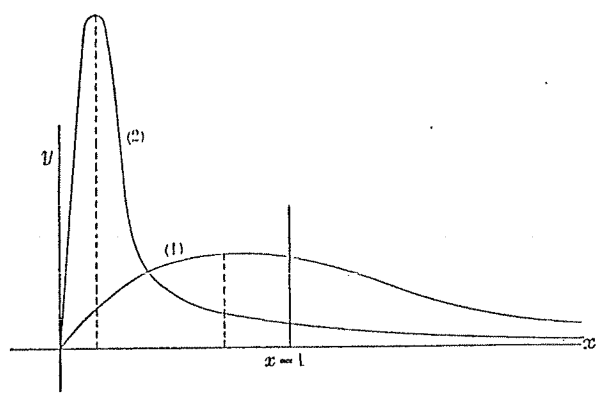
In this example, the history has stayed in the background, but by the time I had finished, the students were anxious to have some details. Since this whole issue has been extensively and hotly debated in the literature over many years, I shall refrain from giving the historical details here. Many of them are in Lakatos's book. For a current entry into the literature, see Laugwitz.[3] Nonetheless, I must end with a historical point. In 1826, a mathematician wrote "... it seems to me that the theorem admits of exceptions" and then provided the first counterexample, the same counterexample that my students had done for homework. The mathematician was the Norwegian, Niels Henrik Abel (1802–1829).[4]

Pictures

The presentation of this material is greatly enhanced today by the use of graphers. The first to adopt the graphical approach was William Fogg Osgood (1864–1943), who, after receiving his Ph.D. under the influence of Felix Klein at Erlangen in 1890 taught at Harvard from 1890 to 1933.[5] The first example that Osgood considers is the sequence of functions

$$s_n = nxe^{-nx^2}$$

where to obtain the graph of the general curve (2), "it is sufficient to divide the abscissas and multiply the ordinates of (1) by $\sqrt{n}$:[6]



This sequence of functions is not uniformly convergent, as is immediately evident from the diagrams.

> For draw the curves $y = f(x)+\epsilon$, $y-f(x)-\epsilon$. Then it is clear that m cannot be taken so large that the approximation curve $y = s_n(x)$ will lie wholly within the belt thus marked off.[7]

[3] Detlef Laugwitz, "Infinitely small quantities in Cauchy's textbooks," *Historia Mathematica*, 14 (1987) 258–274.

[4] Niels Henrik Abel, "Untersuchungen über die Reihe: $1+\frac{m}{1}x+\frac{m\cdot(m-1)}{1\cdot 2}x^2+\frac{m\cdot(m-1)\cdot(m-2)}{1\cdot 2\cdot 3}x^3+\cdots$ u. s. w.," *Journal für die reine und angewandte Mathematik*, 1(1826), 311–339. Reprinted in French translation in Abel's *Œuvres complètes* (1881), 1, 219–250. Also in Ostwald's *Klassiker*, #71.

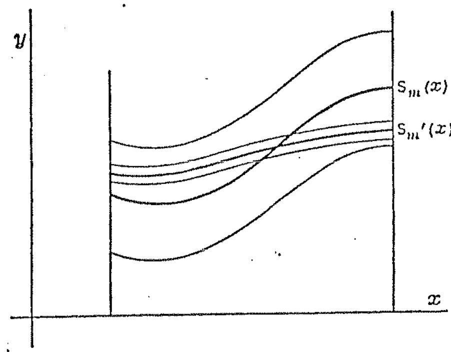
[5] For information on Osgood see J. L. Walsh, "William Fogg Osgood," pp. 79–85 in *A Century of Mathematics in America, Part II*, edited by Peter Duren et al., AMS, 1989 and *A Semicentenial History of the American Mathematical Society, 1888–1938* by R. C. Archibald, AMS 1938, pp. 153–158, which contains a bibliography of his works.

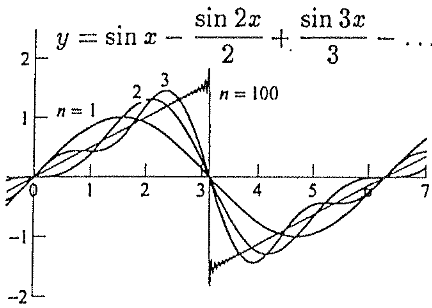
[6] W. F. Osgood, "A geometrical method for the treatment of uniform convergence and certain double limits," *Bulletin of the American Mathematical Society*, series 2, volume 3 (1897), pp. 59–86.

[7] Osgood, *op cit.*, p. 66. Matthias Kawski of Arizona State University pointed out on the calc-reform email list (24 May 1996) that there is a

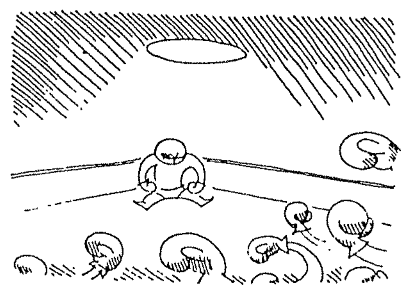
Exercises:

1. Work out the details of Abel's counterexample to Cauchy's "Theorem." This is a fairly hard example to work out in detail, but it is easy to convince yourself by the use of a grapher:[8]





2. Construct a problem about Cauchy's Famous Wrong Theorem that shows the Max can be infinite.[9]



PROBLEMS

Problems worthy
of attack
prove their worth
by hitting back.

---

beautiful graphing calculator view of the difference between uniform continuity and pointwise continuity (and which is applicable, mutatis mutandis, to the corresponding notions of convergence). Pictorically, a function is continuous at a point if given a viewing window of any height (2ϵ) centered at that point, one can always find a width (2δ) such that the graph exits the window only through the sides and not through the top and bottom. This is nice for classroom exploration because students will naturally choose different heights for their windows. Now if one traces the graph, keeping the same window size, and if the graph never exits the top or bottom of the window, then the function is uniformly continouous over the interval that you have traced. Students easily observe that continuity implies uniform continuity on closed bounded intervals.

[8] This picture is from E. Hairer and G. Wanner, *Analysis by Its History*, Springer, 1996, p. 212. This book contains a great deal of interesting historical material.

[9] An earlier version of this note appears in my paper "My favorite ways of using history in teaching calculus," pp. 123–134 in *Learn From the Masters*, edited by Frank Swetz et alia, MAA, 1995.