Notes from the Mathematical Underground

by Alain Schremmer.
Mathematics Department, Community College of Philadelphia,
1700 Spring Garden Street, Philadelphia, PA 19130
SchremmerA @aol.com

The opinions expressed are those of the author, and should not be construed as representing the position of AMATYC, its officers, or anyone else.

Many reasons have been adduced for the 50% failure rate in first semester calculus—lack of …, lack of …, lack of …, but I would submit that not only must a theory in which limits are "the central notion" necessarily fail, *globally*, to impart the differential calculus of functions with a "story line"—as I have intimated before and will discuss at some length in the Fall issue—but, *locally*, it also makes it brutal, right from the start. To dispense with proofs, as is now usually the case, leaves the students no recourse other than to memorize. So, in order to reinstate proofs as something helpful to the students, it is *limits* that ought to be dispensed with. Consider, for instance, the following

THEOREM. *If f is differentiable at* x_0, *then f is continuous at* x_0.

whose Bolzano-Cauchy-Weierstrass (BCW) proof runs something as follows

PROOF. To show that $\lim_{h\to 0} f(x_0+h) = f(x_0)$, we will show that $\lim_{h\to 0} \left[f(x_0+h) - f(x_0) \right] = 0$.

$$\begin{aligned}\lim_{h\to 0} \left[f(x_0+h) - f(x_0) \right] &= \lim_{h\to 0} \left[\frac{f(x_0+h) - f(x_0)}{h} \cdot h \right] \\ &= \lim_{h\to 0} \left[\frac{f(x_0+h) - f(x_0)}{h} \right] \cdot \lim_{h\to 0} h \\ &= f'(x_0) \cdot 0 = 0\end{aligned}$$

END OF PROOF

Now, how can we expect "just plain folks" to *come up* with that? Not being already adept at reducing binary relations to unary ones by way of difference operations, as in "$x > y$ iff $x-y$ is positive" or "*p* entails *q* iff *p*   *q* is tautological", how can they be expected to think of replacing $\lim_{h\to 0} f(x_0+h) = f(x_0)$ by $\lim_{h\to 0} \left[f(x_0+h) - f(x_0) \right] = 0$? Or of dividing and multiplying by h? Worst of all though, because it does not argue straightforward from differentiability to continuity, the proof does not *convince* them: *At this stage*, proofs ought to have some *inevitability*—some logic?—to them. So, indeed, why bother?

On the other hand, since constant functions are the simplest functions with no jump in output and affine functions the simplest ones with no jump in rate of change, why not present continuous functions as locally approximately constant and differentiable functions as locally approximately affine? The proof of the above theorem is now straightforward: In "words", PROOF. Since f is approximately affine and affine functions are approximately constant, f is approximately constant. END OF PROOF

or, in "approximate" algebra, with (…) read "a little bit",

PROOF. Since f is differentiable, $f(x_0+h) = A_0 + A_1h + (\dots) = A_0 + (\dots)$ which says that f is continuous. END OF PROOF

To make sense, the calculus of (…) needs only be based on heuristic considerations of powers of $\frac{1}{10}$: e. g. h^3 is negligible relative to h^2 just as $\frac{1}{1000}$ is negligible relative to $\frac{1}{100}$ and is thus quite within the reach of just plain folks.

But what we are really doing is to define the *order of magnitude* of a function and we do it in exactly the same manner as with numbers: Just as powers of 10 are ordered by the exponent—e.g. when $p>q$, $\left(\frac{1}{10}\right)^p$ is smaller than $\left(\frac{1}{10}\right)^q$, so are power functions near 0 (and ∞): When $p>q$, h^p **approaches 0 faster** than h^q in that Graph $[h^p]$ lies under Graph $[h^q]$. More precisely, when $p>q$, there is for any λ a neighborhood of 0 in which Graph $[h^p]$ lies under Graph $[\lambda h^q]$. Then, we can **gauge** a function f by its position relative to $[h^q]$ and write $\boldsymbol{f(h) = \mathrm{o}[h^q]}$ to mean that $f(h)$ approaches 0 faster than h^q—Graph $f(h)$ lies under Graph $[h^q]$. In BCW, we would say that $\lim_{h\to 0} \frac{f(h)}{h^q} = 0$. For example, if $p>q$, $h^p = \mathrm{o}[h^q]$. In particular, $f(h) = \mathrm{o}[h^0]$ means that $\lim_{h\to 0} f(h) = 0$. Writing o[1] for $\mathrm{o}[h^0]$, note that o[1] behaves like 0 and that it is all we need as $\mathrm{o}[h^p] = h^p \bullet \mathrm{o}[1]$:

$$(\mathrm{o}[1])^\alpha = \mathrm{o}[1] \text{ for any positive rational } \alpha,$$

$$(1+\mathrm{o}[1])^\alpha = 1 + \mathrm{o}[1] \text{ for any integer } \alpha.$$

However, using $\mathrm{o}[h^p]$ is much more intuitive and requires essentially only

$$\text{If } p>q, \text{ then } \alpha\,\mathrm{o}[h^p] + \beta\,\mathrm{o}[h^q] = \mathrm{o}[h^q] \text{ for all real } \alpha, \beta$$

We can either postulate these identities or derive them from BCW but, even then, this would be the last we would need of limits. Note that this allows us to adapt the degree of "rigor" of the proof to the "mathematical maturity" of the students.

Once we can **expand** a function f near any given point x_0, that is once we can express $f(x_0+h)$ as a linear combination of gauge functions plus a **remainder** $R_0(h)$ of which all we need to know is that $R_0(h) = \mathrm{o}[h^n]$, the *local* analysis of f is trivial: For instance, we have that f is **continuous** at x_0 iff $f(x_0+h) = A_0 + \mathrm{o}[1]$ and f is **differentiable** at x_0 iff $f(x_0+h) = A_0 + A_1h + \mathrm{o}[h^1]$ whence the "rigorous"

PROOF. To show that if $f(x_0+h) = A_0 + A_1h + \mathrm{o}[h]$ then $f(x_0+h) = A_0 + \mathrm{o}[1]$ we must show that $A_1h + \mathrm{o}[h] = \mathrm{o}[1]$. Since $A_1h = \mathrm{o}[1]$, we have $A_1h + \mathrm{o}[h] = \mathrm{o}[1] + \mathrm{o}[h] = \mathrm{o}[1]$. END OF PROOF

Since letting $h = 0$ gives $A_0 = f(x_0)$, define the **derivative** of f as the function f' such that $f'(x_0) = A_1$. *Then,* you can reasonably expect students to *come up* with the proofs of the "derivative rules" since $[f*g]'(x_0)$ is the coefficient of h in $[f*g](x_0+h) = [f(x_0)+f'(x_0)h+o[h^1]] * [g(x_0)+g'(x_0)h+o[h^1]]$. If $*$ is defined pointwise, the computations are straightforward; if $*$ is composition, we still have

$$\begin{aligned}[g\circ f](x_0+h) &= g(f(x_0+h)) = g(f(x_0)+f'(x_0)h+o[h^1]) \\ &= g(f(x_0)) + g'(x_0)[f'(x_0)h+o[h^1]] + o[h^1] \\ &= g(f(x_0)) + g'(x_0)f'(x_0)h + o[h^1].\end{aligned}$$

Of course, you will be glad to note that $f'(x_0) = \dfrac{f(x_0+h)-f(x_0)}{h} + o[1]$ and that we obviously do not need L'Hôpital's Rule.

How we expand a function depends of course on its nature. Starting with the positive-power functions as gauge functions, we index polynomial functions with Low d°, the exponent of the *lowest* power term to get a comparison theorem near 0: When Low d°P > Low d°Q, P approaches 0 faster than Q. As a result, $P(h)$ is approximated by any truncation. Ditto near infinity with High d°. We get $P(x_0+h)$ from the binomial theorem and find that, *locally*, polynomial functions are approximately polynomial everywhere!

Gauging rational functions requires division in *ascending* powers, except at infinity where powers must *descend*. Since positive-power functions are not closed for division, rational functions are approximately polynomial only *almost* everywhere. To gauge rational functions near their poles—and near infinity when High d° < 0, we must include the negative-power functions among the gauges.

Beyond that, even though functions cannot be defined by a rule giving the output in terms of the input but, again like numbers, only as solution of—functional—equations, the finite local theory remains elementary because little ohs now do what d° did for polynomial and rational functions: To say of a function that it is $o[h^q]$ is the same as saying of a polynomial function that its Low d° is q: If $f(h) = o[h^p]$ and $g(h) = o[h^q]$ with $p>q$, then $f(h)$ approaches 0 faster than $g(h)$.

For example, ROOT(x) is defined as solution of the equation $f^2(x) = x$ which gives the approximate solution $\text{ROOT}_2(x_0+h) = \sqrt{x_0} + \dfrac{1}{2\sqrt{x_0}}h - \dfrac{1}{8x_0\sqrt{x_0}}h^2$ by the method of undetermined coefficients and, with a leap of faith, that it is an approximation of the exact solution: $\text{ROOT}(x_0+h) = \text{ROOT}_2(x_0+h) + o[h^2]$. So, ROOT is approximately polynomial everywhere except at the origin so that we now must include the fractional-power functions among the gauges.

Define EXP(x) as solution of the initial value problem $f'(x) = f(x)$, $f(0) = 1$. We can expect students to *come up* with $\text{EXP}_n(h) = 1 + h + \dfrac{h^2}{2} + \dfrac{h^3}{3!} + \ldots + \dfrac{h^n}{n!}$ and, when x_0 = is small, with the addition formula: $\text{EXP}_n(x_0+h) = \text{EXP}_n(x_0) \bullet \text{EXP}_n(h) + o[h^n]$ from which a leap of faith gives $\text{EXP}(x_1+x_2) = \text{EXP}(x_1) \bullet \text{EXP}(x_2)$. In particular, $\text{EXP}(n) = \text{EXP}(1)^n$ and, with another leap, $\text{EXP}(x) = e^x$. EXP is approximately polynomial everywhere except at infinity where, since it beats any power function, EXP cannot be gauged. So, it too must be included among the gauges.

Defining SIN(x) as solution of $f''(x) = -f(x)$, $f(0) = 0$, $f'(0) = 1$, we can again expect students to *come up* with, say, $\text{SIN}_3 h = h - \frac{h^3}{3!}$ and $\text{COS}_3 h = 1 - \frac{h^2}{2}$ along with their addition formulas: For instance, $\text{COS}_2(h+k)$ involves $h \bullet k$ which can only come from $\text{SIN}_2 h \bullet \text{SIN}_2 k$. Also, note that $\frac{\text{SIN}_n h}{h} = 1 + \text{o}[1]$! Etc.

I can already hear the shrieks of anguish and terror: "*Mathematics is not a matter of faith*", "*You are evidently and sadly misleading your students!*" Even if, between leaps, students can *prove* things *for and by themselves*? Even if the leaps are clearly indicated as such and if they can, in fact, be fruitfully discussed?

In particular, many of theses gaps occur in situations where we would need to know that some infinite processes reduces in fact to a finite one. (In other words, situations that bring up the need for *compactness*.) For example, say we wanted to compute EXP(x) for some finite x. Start with a small x_0 and compute $\text{EXP}(x_0) = 1 + x_0 + \frac{x_0^2}{2} + \ldots + \frac{x_0^{n_0}}{n_0!}$ for some n_0. Now let $x_1 = x_0$ plus some h_1 and compute $\text{EXP}(x_1) = \text{EXP}(x_0) \bullet \left\{1 + h_1 + \frac{h_1^2}{2} + \frac{h_1^3}{3!} + \ldots + \frac{h_1^{n_1}}{n_1!}\right\}$ for some n_1. Then let $x_2 = x_1 + h_2$ and compute EXP(x_2). Etc. The question is whether the x_i will reach x and, if so, what the error on EXP(x) will be. We can try to take larger and larger n_i but, even so, we may still need to take smaller and smaller h_i to control the size of the error but with the result that the x_i may approach x by diminishing increments and not even get close to x.

Similarly, our definition of continuity at x_0 shows very clearly where the difficulty is in proving theorems such as *a function that is continuous on a closed bounded interval* $[a, b]$ *is itself bounded*. We have that x_0 $[a, b]$, $f(x_0+h) = f(x_0) + \text{o}(1)$. Suppose h is in a neighborhood of 0 whose size depends on x_0, say such that $\text{o}(1) < 0.1$. If we could cover $[a,b]$ with some *finite* number N of such intervals, then $|f(x)-f(a)|$ would be bounded by $\frac{N}{10}$ and the theorem proved.

What then are the little ohs' … limitations? The obvious ones are actually those of the gauges. For instance, just as, near $+\infty$, EXP is too steep and LN too flat to be gauged by power functions, $\text{EXP}\left(-\frac{1}{x^2}\right)$ is too flat near 0; but all this means is that we need finer gauges. However, there are other ways in which functions can misbehave. For instance $f(x) = \cos x + x^3 \sin \frac{1}{x}$ is approximated near 0 by $1 - \frac{x^2}{2}$ while $f''(0)$ does not exist in BCW. But then, this is rather an advantage if the concavity of f at 0 is what we are after! The real challenge is *global* analysis. However, as I showed in a previous column, even there, little ohs allow for a much clearer analysis of the situation than the usual mechanical reliance on critical points. So:

"Little ohs anyone?"
"You may be right but I already made up my mind!"

Friends are deriding us for not submitting (Mattei & Schremmer, 1996) to Messrs. Wiley and Sons! (Rumor has it they are fishing for a *new* calculus author.)

COMPUTERS & DATA ANALYSIS DEPARTMENT. Speaking of the Wileys. In their indefatigable effort to promote the Harvard Calculus, they publish "site reports" in one of which, a Fall 1997 report on the Multivariable Calculus installment of the Magnum Opus, we can read: "*One way in which Multivariable Calculus could be enhanced, from my perspective, would be to make available many problems that are realistically impossible without a computer.*" My undercover agent at Wiley tells me that they are contemplating problems such as:

> According to the Census Bureau, the wealthiest 5 percent of American households increased their share of the nation's income from 15.6 percent in 1981 to 21.4 percent in 1996. The same top 5 percent holds an even larger share of the nation's accumulated wealth, accounting for 60 percent of all the net worth of the nation. (Molly Ivins in the Progressive, December 1997.)
>
> 1. Going on the Internet, get the full Census Bureau data as well as that for the Federal Reserve Board data quoted in the Fall 1997 issue of these Notes.
>
> 2. Using a mathematical modeling program, plot the relevant functions and, in particular, find their limit as t approaches infinity.
>
> 3. Using your brain, discuss the implications for "just plain folks" in general and for their mathematical education in particular.

More seriously, there appeared in a monthly—not *the* Monthly of course, *they* wouldn't be caught dead publishing that sort of thing—an article that should be required reading for anybody even dreaming of bringing the computer to her/his classroom (Oppenheimer, 1997). Here is how it starts:

> In 1922 Thomas Edison predicted that "the motion picture is destined to revolutionize our educational system and … in a few years it will supplant largely, if not entirely, the use of textbooks. " Twenty-three years later, in 1945, William Levenson, the director of the Cleveland public school's radio station, claimed that "the time may come when a portable radio receiver will be as common in the classroom as is the blackboard." Forty years after that, the noted psychologist B. F. Skinner, referring to the first days of his "teaching machines," in the late 1950s and early 1960s, wrote, "I was soon saying that, with the help of teaching machines and programmed instruction, students could learn twice as much in the same time and with the same effort as in a standard classroom.

And then there was television and then there were calculators and now there is …

References

Mattei, F., & Schremmer, A. (1996). *Elements of Differential Calculus: A Lagrangian Introduction to the Theory of Functions*. (5.3 ed.). Rutledge, PA: Courtland Press.

Spring 98

Oppenheimer, T. (1997, July). The Computer Delusion. *The Atlantic Monthly.* 45-62.