Calculus Reform—For the Millions

David Mumford

About twenty years ago I was part of a group of professors from many fields who met once a month for dinner and an after-dinner talk given by one of the members. It was entertaining to hear glimpses of legal issues, historical problems, discussions of what freedom meant in different cultures. But then I had to give one of these talks! I was working on algebraic geometry then, and I tried to figure out what I could say that would hold my colleagues' attention while digesting a large meal. In the end I decided to stick to a rather anecdotal level but to inject one bit of real math. I thought I would try to explain to them the first mathematical formula that I had seen in school which totally bewildered me. This was $e^{i\pi} = -1$. It seemed to me that here was a nugget of real math and maybe it could be explained. Here is how I tried.

I got e into the picture by discussing a savings bank which pays 100% interest, and convinced them that in a year they would get more than \$2 for each \$1 invested. It was not hard to convince them they would get between \$2.50 and \$3 per dollar invested, and we could define e to be their balance after a year. Then I needed i. Most people have heard of i, and I just described it as part of a game invented by mathematicians to get enough numbers so every equation has solutions: starting with the one new rule that multiplying i by i you get -1, you get all the numbers $a + ib$ and their arithmetic. Next, imagine you go to the neighborhood savings bank, and it is running a special promotion with a new account which pays *imaginary* interest at the rate of 100%. The audience immediately sees that you get imaginary interest building up and that the interest on the interest is decreasing your total of *real* dollars. You run them through a few more numbers, and they see that their real funds will go to 0, while their imaginary funds build up to about $1.i$; then they go into real debt, next imaginary debt, and finally get their real \$1 back, which they immediately withdraw! A little picture in the complex plane convinces them that this will take 2π years, while after π years they were in debt \$1 : voilà, $e^{i\pi} = -1$!

What is the point of this struggle to communicate some tiny bit of math? For me, the lesson was that I think my audience got a bit of honest math from this and that what they struggled to learn consisted of some numerical fiddling, some geometry (the circle showing the evolution of your balance in the complex plane), and some thinking about the rules which underlie arithmetic and exponentiation. I do not think I said anything mathematically dishonest, yet I certainly gave *no proof of anything*. I think this is the same approach as that taken by many calculus reform texts, and it is exactly the philosophy of the Gleason-Hallett Calculus Consortium.

The problem of communicating comes up in many situations other than after-dinner speeches. For nearly fifteen years I have been doing applied mathematics, and I have to talk especially to biologists, psychologists, and engineers. The same rules seem to apply. If I mention L^p, they have me pegged as "one of them".

David Mumford is University Professor, Division of Applied Mathematics at Brown University. His e-mail address is David_Mumford@Brown.edu.

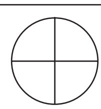
Of course, the tolerance level varies. Some engineers have been rather thoroughly mathematicized: in control theory everything is done in multiple Banach spaces. But I know a psychologist who hates math yet understands absolutely correctly the meaning of robust statistics. For pure mathematicians and statisticians, robust statistics refers to statistics that work when the variables being measured are not normally distributed and still give you good estimates of things like the mean and variance of the variable. In real life I think it is fair to say that nothing is ever normally distributed because there are always "outliers", exceptional cases which are off the scale. My psychologist friend spent much of his time measuring reaction times and would average over his subjects to get a mean. When I did the same and got terrible results, he said, "Maybe you forgot to throw out the slowest 3 percent?" The point is that about 3 percent of the time, his subjects' minds wandered, and he got absurdly slow responses. Is this math? In fact, yes: there is a very substantial body of theory on what "α-trimmed means" do for you with unknown distributions. My friend has an excellent intuitive grasp of this without knowing any of these theorems.

Often a picture is what facilitates communication. When you were in grade school, you might have been puzzled, as I was, when asked to accept the formula $\frac{1}{\frac{1}{a}} = a$. Of course, modern books "prove" this, more or less, manipulating the axioms in the usual way. But does it not become just as clear from a picture which compares:

((••)(••)(••))

How many pairs is 6? **Answer:** 3

As a formula: 6/2 = 3



How many quarter pies in a whole pie? **Answer:** 4

As a formula: 1/(1/4) = 4

An example, a picture, an explanation is presented. Would it be better to present a proof?

I learned calculus during high school when I stumbled across a great classic of the pedagogical literature: Lancelot Hogben's *Mathematics for the Million*. Hogben explains the essence of calculus, including differentiation, integration (both with many examples up through trig functions), the fundamental theorem, and multivariable integration through Archimedes' great achievement, calculating the volume of a sphere. He does this *in fifty pages*! Was he successful? Well, the book went through four editions over more than thirty years, apparently being read by literally millions. How did he do this? Here is how he introduces the derivative:

> If the points p and q in the course (of a cyclist) are very close together, … the curved line joining them is difficult to distinguish from a straight line, and the pointer of the speedometer will not shift appreciably during the interval representing the difference between the x co-ordinates of q and p. When p and q are very close together, so that we cannot distinguish them, the line passing through them becomes the tangent at the point $p = q$, and the gradient of this line corresponds with the speedometer-reading at the instant represented by the x co-ordinate of p.
>
> ....
>
> The tangent method is equivalent to taking two points with x co-ordinates x_p and $(x_p + \Delta x)$ and y co-ordinates y_p and $(y_p + \Delta y)$ so close together that Δx and Δy are too small to measure. The gradient is $\frac{\Delta y}{\Delta x} = \tan a$. When Δy and Δx are immeasurably small, we write the ratio (pronounced as dee-wy-by-dee-eks) $\frac{dy}{dx}$.
>
> (pp. 521–522, 3rd edition)

Hogben was a genius at putting things in plain English. He used Δ, but no ϵ or δ. He was also very clear about the need to explain calculus in down-to-earth terms. He railed against Newton himself:

> The intellectual leaders in the Newtonian period did not realize that every intellectual advance raises a constructive problem in education. Newton himself devoted much of his energy to devising long-winded demonstrations in Euclidean geometry instead of trying to make his own methods intelligible to his contemporaries. One result of this was that conspicuous progress in Newtonian mechanics did not take place in his own country during the century which followed the publication of the *Principia*.
>
> (op. cit., p. 567)

A striking contrast for me was the curriculum which my oldest son encountered learning Euclidean geometry in a Paris high school in 1976. Unfortunately, I no longer have the textbook, but the following is close to the definition presented there of a "Euclidean line": a *Euclidean line* is an ordered pair $\{X, \Phi\}$. Its first member is a set X whose elements will be called "points". Its second member is a set Φ of bijections between X and the real line $\mathbf{R}$ which satisfies two axioms. First, for all $\phi, \psi \in \Phi$, the composition $f = \phi \circ \psi^{-1}$ is a map from $\mathbf{R}$ to itself of the form $f(x) = \pm x + a$ for some real number a. Second, for any $\phi \in \Phi$ and any map $f(x) = \pm x + a$, there is a unique $\psi \in \Phi$ such that $\psi = f \circ \phi$.

I believe the same concept was circulating at the time in research circles under the name "torseur". It seemed to me at the time a bizarre way to prepare the next generation of educated Frenchmen. But perhaps Hogben is right and Euclidean geometry has this effect on the abstract thinker.

Watching my children move through the mathematical curriculum of elementary and high schools has been very instructive for me. For instance, I believe there is *no universal best way to teach mathematics* which applies to all the basic skills. It is very tempting to adopt some pedagogical theory or intellectual standpoint and convince yourself that this is the yellow brick road leading to understanding. I am not convinced that the experts who study pedagogy in mathematics have a deeper insight into what works than most concerned parents.

To illustrate, at one extreme I suggest there are some essential topics which *must* be memorized. The multiplication tables are the prime example. People with numerical gifts see the patterns in the tables and use these to learn them faster, but by and large a formula like $7 \times 8 = 56$ has to be memorized (one of my kids learned it because *Creature Double Feature* was on from 7 to 8, Channel 56).

There is a large chorus of people who rail against teaching calculus by "cookbook" methods. But my gut feeling is that some topics needed by everyone in a numerical profession are learned fastest by taking them purely as the rules of a game prescribed by the inventor of that game. Solving equations in algebra is a prime example. If these are taught like a board game, with rigid rules about when you can move a piece from one square to another, they are not much harder than checkers, say. For me, $b^2 - 4ac$ still has the flavor of a memorized icon. One hopes the meaning will come with practice and application. But drilling in cookbook methods seems a reasonable method for bootstrapping the skills of high school students to the level where they can begin to deal symbolically with algebraic relationships. We drill students in conjugating French verbs, so why not in algebra?

But everyone agrees that this approach of memorizing and game playing has real limits. Learning to correctly convert among fractions, decimals, and percents was such a case for some of my children. I do not think this can be learned either by memorization or by pretending it is a bizarre game required to pass tests. It is also an essential skill in later life in thinking about budgets, inflation, savings, and using recipes. What it seems to require is an understanding of what it means based on many simple but real examples: converting dollars to cents, converting proportions in a recipe to numbers of ounces, etc.

Focusing on calculus reform, let us distinguish three types of pedagogical methods. One is memorization and drill. Because of the examples just given, I would defend the proposition that these have their place; they have certainly been used to attain high standardized test scores in calculus as well as in arithmetic and algebra. But we all know that, although often effective for short-term results, this approach has its limits if the concept is subtle or will not be practiced regularly. Another is the use of many examples, numerical and visual and based on things already familiar to the student. This aims at the gradual solidification in the student's mind of an intuitive gut feeling of the meaning of the concept. This is what I think is needed for decimals and fractions. It is what Hogben did so successfully in his classic book and what the Gleason-Hallett reform text aims to do. The third is the presentation of the underlying logic of the theory, making an airtight legal case that such and such and nothing else must be true. All professional mathematicians are in love with this, I among them. The "new math" attempted to bring in logical arguments at the very early stage of basic arithmetic—for instance, by *proving* rules like $\frac{1}{\frac{1}{x}} = x$. The French definition of a Euclidean line is an extreme example of how formal definitions are introduced so that complete proofs can be given.

A very interesting point relating to the use of formal definitions has been raised by Saunders Mac Lane and other critics of the Gleason-Hallett text. They object to the definition of a continuous function—"the closer x gets to a, the closer $f(x)$ gets to $f(a)$" in this book—raising the example of $x \sin(1/x)$. The problem here is that English syntax is notoriously ambiguous in common usage, and the intended meaning is often inferred from common sense rather than from any general syntactic or semantic rules. To give an example of ambiguities in normal English usage, an example much discussed by linguists like Montague is "Three lighthouse keepers saw three ships". In this sentence, the problem is that each lighthouse keeper might have seen a single one out of a set of three distinct ships, they might all have seen the same three ships, or maybe they each saw a ship and no one is sure whether they are the same or different ships. For the sentence above which purports to define continuity, I suggest you ask a mathematically naive friend whether they find anything odd about the sentences: "A clock pendulum is slowing down from friction. As it does so, it gets closer and closer to the vertical position." Or ask whether the assertion that "runner x is getting closer and closer to a new world's record in the 100-meter" implies that x never has a bad day? I certainly agree that a footnote clarifying the Gleason-Hallett definition to say that $f(x)$ need not *go straight* to $f(a)$, but may wobble on the way, is appropriate. But whether the Gleason-Hallett definition is correct as it stands is not a well-posed question: virtually the only sentences in English with an unambiguous interpretation (not requiring the use of common sense by the reader) are those written in mathematical jargon.

If one says instead, "For any $\epsilon > 0$, there is a $\delta > 0$ such that whenever $|x - a| < \delta$, then $|f(x) - f(a)| < \epsilon$," what happens to most students? First off, since Greek letters and complex English syntax ("for any … there is … such that whenever … then …") are used, the student is convinced that something very complicated must be going on. What is worse, even if you give the simple description of the meaning afterwards, the student will be sure that something more complex is going on or else why did you put it in such an opaque way? I think it is impossible to explain to most students that we prefer the complex syntax of the $\epsilon - \delta$ definition because we have crafted it precisely to squeeze out all the ambiguity of normal English. The important question is: do *most* people learn a new concept most efficiently by being exposed to elegant definitions of this sort? For instance, if your neighbor happens to ask you what you are teaching, and this happens to be calculus, how do you explain the derivative over the fence to him/her?

Wu makes the case in the December *Notices* that even if we admit full rigor is inappropriate for nonhonors students, we should at least aim to train them in making logical deductions. Why? Pure mathematics is the only discipline in which proofs are deemed so basic to knowing the truth. Especially since the advent of Ed Witten on the mathematical scene, we have realized all too clearly how physicists (and mathematical physicists) can often get at the deepest sort of truths without paying any heed to rigorous proofs. He is only the most visible example, as all of modern physics is built on "derivations" which are combinations of heuristics, calculations, and occasional precise arguments. In other sciences, such as chemistry and biology, logical deduction has virtually no place, because the systems being studied are too complicated to allow one to prove anything rigorously: what scientists do is to argue that such and such is the most likely explanation, based on data, analogies with other systems, and appeal to a shared Bayesian model of what one expects this kind of system will do. If scientists use logic so rarely, this is even more true for the rest of the educated public. In political discourse we not only fail to see logic used, we do not even see people using numbers coherently to quantify the issues. If we as a community want to take up an educational cause, maybe we would do better to try to get a larger group of people to believe that numbers can help them understand the world around them.

Applied mathematicians vary hugely on this scale, but many are much more interested in testing a model by simulations than by the much harder rigorous proof. Since all models are incomplete on a fundamental level because they isolate only a few of the complexities of nature, this test by simulation is often the really crucial one. How many years have gone by since Lorenz simulated his three-dimensional dynamical system without anyone being able to rigorously analyze his system? This is typical of nonlinear differential equations. Hodgkin and Huxley won the Nobel prize for their family of PDEs which model conduction on nerve axons. The clincher in their work was their computer simulation which correctly predicted the speed of conduction to within 10 percent. But only toy versions of their equations have been proven to actually produce stable traveling waves. In my own experience in computer vision, a nonlinear parabolic equation was proposed to enhance images. The equation was patently ill-posed, but works extremely well in simulations! The message for pure mathematicians is not to throw it out, but to find a well-posed problem which somehow captures the same significant behavior. The moral in all these cases is that the lack of proofs or even of well-posed models does not inhibit good applied mathematical modeling.

In the nineteenth century there was no clear division between pure and applied mathematics, and people like Riemann went back and forth between the two areas. He was quite satisfied with his use of the Dirichlet principle to prove the existence of solutions of various PDEs, although a rigorous justification of this argument took decades.

I believe I am on firm ground in stating that only lawyers[1] love proofs the way we do. The only colleague in the dinner club mentioned at the be-

[1] *A referee kindly pointed out that religious scholars form another group which loves proofs too!*

ginning of this article who afterwards wanted to borrow a math book from me was in fact a professor in the law school. But a lawyer's idea of formal argument is so convoluted that it has resisted all attempts by computer scientists working in artificial intelligence to formalize it. A widespread disinterest in proofs is surely one reason why pure mathematics is the most isolated of the sciences. If we could give up our obsession with always being so precise and communicate more loosely what we are doing, we might break into the *New York Times* "Science" section more often. Only when a romantic hero like Erdös, a mathematician's mathematician who maintained a master list of "God's proofs", comes along can we break this barrier. We are intoxicated with the depth and subtlety of things like nonstandard four-space and the proof of the Fermat conjecture. But as a profession we are not very successful in communicating this beauty.

In summary, we have scientists, engineers, economists, and people in the world of affairs in one category—call it P (for "practical"); and we have the twentieth-century community of professional pure mathematicians in another—call it T (for "theorem-loving"). Applied mathematicians, lawyers, and mathematicians of other centuries fall somewhere in between. Many people in group P use calculus. A calculus course is often the last interaction between these two worlds. I would guess something like 99 percent of our students in these courses are not going to join category T. So calculus is our big chance to talk to the other world, P. We have two strategies open to us. One is to use this opportunity to preach the gospel of logic and reveal the beauty of precise definitions. Two examples that are often given are the rigorous definition of limit with ϵ and δ and the mean value theorem. On the other hand, the calculus reform movement (or some of the heads of this many-headed monster) takes the position that neither of these helps the students in group P understand better what calculus is about or what it is good for. Speaking to my mathematical friends and colleagues in group T to find out why they prefer a rigorous approach, I found that for many of them, their first exposure to this sort of rigor was a defining experience in their lives. I seem to be in the minority in having learned calculus from a pedestrian book like Hogben's. But the question is whether there are very many graduates of calculus in group P who found this exposure to rigor to be equally significant. I doubt it. Are we teaching calculus in the hope that a small percentage of our students will catch our love of rigor, or so that most of our students will emerge with the ability to use calculus in their specialties?