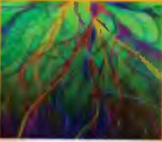


The Mathematician's Brain

DAVID RUELLE

❖ 18 ❖

Mistake!

THE LANDSCAPE OF mathematics has a historical dimension. New theorems are proved, and better tools are devised for handling all kinds of questions. At the same time, the problems that remain unsolved become progressively more difficult. I once had a chance to chat about this changing world with Shiing-shen Chern,[1] one of the great figures of twentieth-century geometry. And he explained to me how, when he started his mathematical career, he read the work of Heinz Hopf on fibrations of spheres[2] and found himself at the frontier of the mathematics of the time; he could start doing his own original work. Now, Hopf's ideas are wonderful but relatively easy to study. At the beginning of the twenty-first century, it is typically much harder to get to the frontier of mathematics. Think of having to master Grothendieck's ideas, among others, if you want to work in algebraic geometry and arithmetic!

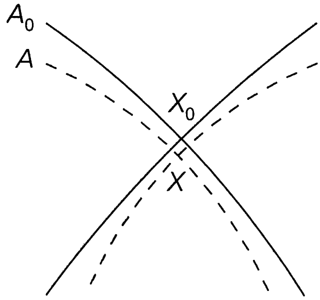
Mathematics does not always become more difficult as time goes by. Sometimes a new technical development provides access to questions that had hitherto been out of reach. Sometimes problems that had not attracted interest become the center of a bright new field of mathematics, with important results relatively easy to obtain. For instance, the arrival of fast computers promoted the study of algorithms and led to basic conceptual developments like the notion of NP completeness and the remarkable proof that primality can be tested in polynomial time.[3]

In general, however, one must admit that mathematics becomes more and more difficult with time. This causes changes in the practice of research. I remember hearing criticism in the 1960s addressed to a mathematician who used results by others without sometimes checking them personally. Because of the inflation of the literature, this checking of earlier results is less and less possible. I heard Pierre Deligne, in the 1970s, stating that the mathematics that interested him was that which he could personally understand in complete detail. This excluded, he said, proofs using computers and extremely long proofs that a single person could not dominate. In fact, however, computer-aided proofs and extremely long proofs have become a normal feature of contemporary mathematics.

Perhaps we are witnessing a decay of the "moral values" of mathematics; Grothendieck said so explicitly. At the same time, however, we see extraordinary successes in the solution of old problems (Fermat's last theorem, the Poincaré conjecture,[4] etc.), and we must admit that contemporary mathematics is in a sense remarkably healthy. Simply, we observe that the nature of human mathematics is changing. And different people adapt to the change in different ways. An example that has led to some controversy arose from William Thurston's work on three-dimensional manifolds. A natural problem of geometry is to *classify* manifolds of a certain kind (this means listing them). The classification of two-dimensional manifolds is well understood. But the study of three-dimensional manifolds is much harder. After considerable work, Thurston came to a good understanding of the subject, of which he gave a broad description, with outlines of proofs. Thurston's program thus laid claim to a big mathematical area but without providing proofs that colleagues could check. In effect, he made it difficult for other mathematicians to work in this area: you don't get much credit for the proof of a theorem that has already been announced, but at the same time you can't use this theorem, because a proof doesn't quite exist yet. A much-discussed paper by Arthur Jaffe and Frank Quinn,[5] mentioning Thurston and others, complained about this evolution of mathematics. As it turns out, Thurston's program has now been largely implemented, but the problem raised by Jaffe and Quinn remains significant for some parts of mathematics.

It is now time for us to look into the mathematical use of computers. Speaking of computers one thinks of long numerical calculations. Are such calculations useful in pure mathematics? Sometimes they are. In fact, Riemann did long numerical calculations by hand to test some ideas and would surely have been pleased if he had had a fast computer at his disposal. Computers have also been of great help in visualizing objects that occur in the theory of dynamical systems.[6] There is thus no doubt that computers can be of use in the *heuristics* of mathematical problems, that is, they make some conjectures plausible and invalidate others. Most mathematicians have no objection to this heuristic role of computers. But the normal use of computers gives only *approximate* numerical results; how can they be used to obtain *rigorous proofs*?

Computers are really rather versatile machines. Let me mention some tasks that they can perform exactly, which can be of use in proving theorems. Exact calculations with integers is the most obvious thing. Computers, however, can also be programmed to do logical operations: checking, for instance, a large number of situations and giving in each case a yes or no answer to some question. This *combinatorial* capability of computers is what was put to use in the proof of the four-color theorem.[7] Computers can also handle real numbers like π or $\sqrt{2}$ exactly, by using *interval arithmetic*. The idea is that, if you know that π is in the interval (3.14159, 3.14160) and $\sqrt{2}$ in the interval (1,41421, 1.41422), you also know that $\pi + \sqrt{2}$ is in the interval (4.55580, 4.55582) *without any error*. Interval arithmetic allows one to perform, with strictly controlled accuracy, all kinds of calculations involving real numbers. Let me sketch an example of how such calculations can be used to prove a theorem. Suppose we know that two (explicitly specified) curves A_0 and B_0 in the plane intersect at a known point X_0, and we want to prove that the (explicitly specified) curves A and B intersect at a point X close to X_0.



This is known to be true under certain conditions (transversality of the intersection of A_0 and B_0, closeness of A to A_0 and B to B_0 in a certain sense) that can be checked numerically. It may be convenient to do the numerical checking with the help of a computer. I have just outlined a computer-aided proof that, under suitable conditions, two curves *A* and *B* have a point of intersection *X*, with an estimate of the distance of *X* to a known point X_0.

As it happens, some theorems of real mathematical interest are of the form just described, but with curves *A* and *B* replaced by manifolds in some infinite-dimensional space. My colleague Oscar Lanford reported once on a theorem of that kind.[8] The contents of the theorem will not concern us here. Rather, we shall look into some technical aspects of how Lanford proved it. The proof is computer aided, which means that it consists of some mathematical preliminaries and then a computer program. This program (or code) uses interval arithmetic to check various inequalities; if these are found to be correct, the theorem is proved. The complications of the problem forced Lanford to write a relatively long program, about 200 pages. The pages consist of two columns; one has the code (in a variant of the C programming language), and the other has explanations of what one is doing. Indeed, long code without explanations is incomprehensible, even to the person who wrote it. And in the present case, since it is a mathematical proof, other people should be able to check it. Oscar Lanford is a very careful person, and he took pains to check that, when the code is fed into the computer, the computer does exactly what it is supposed to do. In this manner—after the computer has agreed with the inequalities in the code—the proof of the theorem is complete.

But Lanford added some remarks that you may find rather disheartening. "I am sure," he said, "that there are some mistakes in the code I wrote. But I am also sure that they can be fixed and that the result is correct." What this means is that in 200 pages of text, there are probably some mistakes. In the present case it might happen that some inequality that needed to be proved was, in fact, not proved! But Lanford believes that he has a sufficiently detailed understanding of the problem at hand and that he could find and prove a similar inequality, sufficient to establish his theorem.

It is good to remember at this point that computer-aided proofs are not completely formalized mathematics (that one could, in principle, trust completely). Computer-aided proofs are part of human mathematics. The problem of avoiding mistakes when you use a computer is different, however, from what it is in "normal" pencil-and-paper mathematics. You can check and flush out some kinds of mistakes present in computer code, but you don't have the intuition that good professional mathematicians have developed about pencil-and-paper proofs.

With proofs getting longer and longer, the problem of mistakes in mathematics is becoming increasingly serious with time, whether or not computers are used. Together with mistakes, I shall here discuss gaps, that is, elements of a proof that are supposed to be easy to see, but aren't. To put it bluntly, the probability that there is no mistake in a proof goes down with its length in an exponential manner (or worse). And a single mistake can kill a proof! It is fortunate that many mistakes, like misspelling a name or putting a wrong date in a reference, are mathematically inconsequential (even if they can make some people quite furious). More serious mistakes can often be fixed too, and we shall later discuss how that happens. We can see how grave the problem of mistakes or gaps can be by looking again at the theorem of classification of finite simple groups. The proof of this theorem covers many thousands of pages, by many authors, and parts of the proof are computer aided. The theorem has been considered as "morally" proved since around 1980, with some parts yet to be written. This means that there were gaps in the proof, but they were not considered serious by specialists. One of those gaps, however, turned out to be serious enough to necessitate another 1,200 pages of proof (in 2004[9]). There are other areas of mathematics that are in a messy state. For instance, speaking of the packing of spheres, Tom Hales wrote, "the subject is littered with faulty arguments and abandoned methods."[10]

Does this mean that mathematics has forgotten its old standards of rigor? That mathematical truth has become a matter of opinion rather than a matter of knowledge? Interesting views on this problem have been expressed by various authors in response to the article of Quinn and Jaffe mentioned earlier.[11] Basically, one may say that good mathematicians working in an area know how reliable the published literature is. Some areas have been scrutinized repeatedly by high-level mathematicians and theorems proved by different methods; such areas can be considered as extremely rigorous. But one must admit that the mathematical literature contains lots of junk, because some people need to publish for career reasons even if they have little interest in what they are doing.

In brief, the old ideals of absolute logical rigor have not been abandoned. But there are forces at work to change the style of mathematics, because highly desirable theorems may require very long proofs or computer-aided proofs. Think, for instance, that if you want to prove a property of simple finite groups, you can do it by checking the property in question on an explicit list of groups. This shows how useful the classification theorem is: it is a new beacon that changes the landscape of mathematics. Of course there are changes also for the human mathematician. Being a mathematician today is not what it was a hundred years ago. Doing mathematics in a hundred years will again be different. Perhaps it will be a less satisfying enterprise than it was in earlier centuries, and perhaps not. But there will be new results, deeper theories. And more of the unknown face of mathematical reality will have come to the light of human understanding.