14

What, If Anything, Is an Experiment in Mathematics?

Jean Paul van Bendegem

1. Introduction

The subject of this paper is an extremely modest inquiry into the nature of mathematical experiments, if such things exist, *as put forward by mathematicians themselves.* I am thus not interested in the no doubt philosophically more rewarding general question: Is there such a thing as a mathematical experiment, and, if so, what kind of experiment is it? Furthermore, 'experiment' is to be understood here in a very naive sense. If one believes that all objects floating freely above the surface of the earth will fall towards the earth, then an experiment to check this statement will consist in taking a material object, holding it above the surface of the earth and letting it loose (to induce the state of floating). An experiment, therefore, has something to do with the manipulation of objects, with setting up processes in the 'real' world and with observing possible outcomes of these processes. Again, this naivete is not meant to avoid the interesting but difficult philosophical conundrums surrounding the notion of experiment. It just so happens that mathematicians themselves seem to have this picture in mind.

Lest I be accused of ignoring (or being ignorant of) a major part of modern philosophy of mathematics, let me emphasize the point made above along a negative path. This paper is *not* about the following:[1]

- The no doubt experimental nature of the learning of mathematics (see, e.g., Ernest, 1991). Children manipulate objects of all kinds and somehow numbers and other mathematical concepts are derived from these experiences. The problem dealt with here is not whether the thing is learned along experimental routes but rather whether the thing learned has something experimental about it.

- The no doubt experimental nature of the growth and development of mathematics. Ever since Imre Lakatos' seminal study "Proofs and Regulations," it must be considered impossible to deny this point of view. I must add here that I am however somewhat troubled by the use or rather abuse of language in this context. We are talking here about the experimental nature of the growth of mathematics. It seems rather odd to draw therefrom the conclusion that mathematics itself is experimental. If such an inference were permissible, then theology too must be considered experimental—a philosophically rather absurd conclusion.
- The nature of certainty in mathematics. In recent years, there has been considerable discussion in mathematical circles about long, difficult and complex proofs. How can we be sure that these proofs are correct (examples include the proof of the four-color theorem and the complete proof of the classification of finite, simple groups)? Are we not forced to give up certainty? Apparently, via such ingenious concepts as 'zero knowledge' proofs and 'transparency' proofs, it is possible to lower the probability that a proof contains an error. Thus we do have to replace the certain nature of mathematics by a probabilistic measure of error. Again, this is a problem not about mathematical statements themselves but about the verification of proofs of these statements. The four-color theorem does state that, with certainty, all planar maps are colorable with no more than four colors. There is no mention at all of probabilities in this statement.

There are probably more uses and interpretations of the notion of 'experiment' in mathematics than the above list offers. But, once again, this paper is written by a philosopher, who is puzzled by the fact that mathematicians *themselves* regularly use the term 'experiment' to describe some of their activities. Are they not supposed to be (open or closet) Platonists? Are we witnessing schizophrenia in mathematics?[2] What is my evidence? All statements I have come across in the literature seem to fall into one of two neatly separated categories. I will, therefore, treat them separately. The first category carries the label "experiment = computation" whereas the second one carries the label "experiment = real-world-experiment."

2. Experiment = Computation

(a) Common folklore relegates mathematics to the theoretical sciences, where information and knowledge is reached by logical steps within an abstract framework, and not from experiments. In fact, mathematical discoveries are more likely to spring from the patiently acquired experience of many specific computations. (Vivaldi, 1989, p. 49)

This author seems to suggest that a large set of computations can be considered as experimental evidence. In fact, he discusses the problem of solving non-linear differential equations for which analytical solutions are not available. Hence, computers are used for calculating numerical solutions giving 'evidence' of chaotic behavior.

> (b) I believe it is best to treat computer-based proofs as scientists in other disciplines treat experimental results. . . . Scientists in other disciplines have standard techniques for handling this problem. They emphasize the importance of independent verification of experimental results. (Lam, 1990, p. 12)
> Maybe we can borrow a further leaf from the experimental physicist's book—they have their particle accelerators, so why shouldn't we dream of mathematical supercomputers? (ibid)

This author has used a computer to show that a specific mathematical structure—a finite projective plane of order 10—does not exist. The computer rather crudely went through all possibilities.

> (c) The example of the four-color theorem may help to clarify the possibilities and the limitations of the methods of pure mathematics and those of computation. It may be that a problem cannot be solved by either of these alone but can be solved by a combination of the two methods. There is a certain parallel in the early history of science. From the time of Plato until the late Middle Ages, mathematical methods were regarded as so superior to experimental methods that experimental physics was not considered acceptable among serious scientists. This severely handicapped the development of certain branches of physics. (Appel & Haken, 1978, p. 179)

This quotation is from the authors who claim to have found a proof of the four-color theorem. As part of their "proof"—I use quotation marks as not all mathematicians are convinced that it is indeed a convincing proof[3]—they used a computer.

What these three examples have in common is that computation(s), with or without the use of a computer, are seen as providing experimental evidence for mathematical statements or problems. It follows that if a computation or series of computations is to count as an experiment, then some similarity has to be found between a computation and an experiment as it is traditionally understood in the sciences. Consider then, first, the following analogy. Let A(n) stand for a number-theoretical statement, n ranging over the natural numbers. Say A(n) is: "A natural number of the form 2n is the sum of two primes" (the well-known Goldbach conjecture, still an open problem). Take its universal closure:

(G) For every n, A(n).

Suppose I now calculate for n = 500, whether or not I can find two primes p and p′ such that p + p′ = 1000. Starting with p = 2, 3, I will soon find the answer 3 + 997 = 1000. Thus, I can write a 'report' of this 'finding':

(G500) A(500)
or, more precisely: 1000 = 3 + 997, thus A(500)

At first sight, it is tempting to see (G500) as confirmational evidence for (G). After all, the more cases we would examine, if all of these turn out to be positive evidence, the more we are justified in believing that (G) probably holds.[4] True, it does not *prove* (G), but that shows how close the analogy runs. No matter how many black ravens you have seen, it does not entitle you to claim *with certainty* that, indeed, all ravens are black.

Plausible though it may seem, a computation cannot qualify as an experiment on this interpretation, for the simple reason that one can *prove* the statement (G500). In comparison, it would be very odd indeed if someone were to claim that he or she can *prove* to have seen or observed a black raven. Whatever definition one tries to formulate for the notion of an experiment, it should include the feature that experiments and observations are information-gathering techniques, not proof techniques. Putting it differently, there is no need for the mathematician investigating (G) to leave the mathematical domain of number theory, to find the statement (G500). In fact, the derivation of (G500) can be carried out on the syntactical level only. There is no need to know what 1000 stands for. Think of Peano Arithmetic and the steps needed to derive 3 + 997 = 1000, or more precisely $0^{(3)} + 0^{(997)} = 0^{(1000)}$, where $0^{(n)}$ stands for 0(the syntactical zero) followed by n applications of the successor function. An experiment, whatever it is, must involve a semantical component. It tells us something about (a model of) the world. In most cases—perhaps making exception for thought experiments[5]—an experiment constitutes an interaction with the world. As such, it is semantically based and cannot be reduced to a purely syntactical game.

As an additional argument, note that the statement 3 + 997 = 1000 itself can be considered a theoretical statement that can be supported by experiments. If we use strokes to represent the numerals, then if we put 997 strokes next to 3 strokes and count them to find 1000 strokes, should that not be considered an experiment? But if so, what is to be counted as an experiment apparently depends on the context. Something can be a theoretical thing from one point of view, yet an experimental thing from another point of view. This, at least, should be considered slightly, if not simply, odd.

My conclusion therefore is that indeed computations are not to be considered as experiments in any sense. If so, this raises a quite different question which I want to say a few things about. Why then do mathematicians (apparently) think of them as such? I think there are at least two reasons. The first has to do with the fact that the computations are (in an increasing way) not executed by humans but by computers. The use of computers introduces a tool in the mathematics trade that was not there before. Not that mathematics has ever been without its tools, but the computer is in all senses of the word a machine. And more importantly, it does things we cannot do. If a computer program goes through a couple of billion cases for (G), then the machine does something that is not accessible to us. It therefore carries with it the extremely strong suggestion that the machine gives us a glimpse of a part of the mathematical universe that is not available to us. In that sense, of course, it seems quite similar to an experiment or observation. The same applies even more strongly to visual graphics. The mathematician does not know precisely what the object she wants to visualize looks like, and then the machine shows the (a) picture. That the similarity causes confusion is quite understandable. In a way, the confusion amounts to the same thing as confusing reality with cyberspace or virtual reality. In fact, there is plenty of evidence that the confusion is already establishing itself.

The second element has to do with the fact that these computations do not, in general, provide any insight. As mathematicians and philosophers seem so fond of the following example, let me take it as a case in point. Suppose you have a 4 × 4 grid. The top-left and bottom-right corner have been removed. One now has a set of dominos such that a single domino covers exactly two squares of the grid. The question to be answered is whether it is possible or not to cover the grid with dominos. There are two ways of dealing with the problem:

- To calculate all possibilities. In the 4 × 4 case, this is quite simple. The reader can check for himself that just two scenarios are possible, and that both of them lead to the same conclusion: one is left with one uncovered square surrounded by covered ones. Hence, no solution is possible. Let me call this the *computational proof.*
- To reason as follows: suppose that the full grid is colored black and white alternatively. Removing the two corners eliminates two squares with the same color. Hence, the remaining grid has not the same number of white and black squares. But any domino will always cover a black and white square. Hence a set of dominos covers the same number of black and white squares. Hence the impossibility is demonstrated. Let me call this the *mathematical proof.*

On the one hand, it is easy to see how tempting it is to make the comparison with the notion of experiment and the notion of theory. If we gather data through experiment or observation, we do not thereby gain (necessarily) any insight into the hidden order or structure present in the data. The data need to be interpreted, they need to be inserted into a model, etc. Likewise, the computational proof does not provide any insight, whereas the mathematical proof does. In fact, the domino-problem is a perfect example to illustrate the point being made here. One sees immediately that the mathematical proof can be easily generalized to a grid of any size. One sees immediately that it is not necessary that the two opposite corners should be removed. Remove any two squares of the same color, and the mathematical proof still holds. Not so for the computational proof. Change any element in the problem data, and one has to start all over again. The mathematician is forced to use a case-by-case method.

On the other hand, of course, it is equally obvious that both ways of handling the problem, are to be considered *proofs and nothing but proofs.* The computational proof *does* solve the problem in a mathematically perfectly acceptable way. Obviously, it is, what one might call, a low quality proof, but a proof nevertheless. Note too that, even if a theoretical well-founded distinction could be made between computational and mathematical proofs,[6] this would still leave us with the problem of mixed proofs. The four-color theorem seems to be an excellent real-life example.

If the distinction between computational and mathematical proofs is difficult to make, then it does not follow that no such distinction can exist. One might ask, is it indeed the case that only mathematical proofs can generate insight? According to some authors (e.g., Steiner, pp. 102–108, Dunham, pp. 212–217), some forms of reasoning to be distinguished from the notion of proof do exist and are used in mathematics. Let me call this a mathematical *argument.* The most famous example is Euler's argument for $\Sigma\ 1/n^2 = \pi^2/6$ (the summation taken over the natural numbers). I will not present the full argument but only its general structure. Euler first reasons about polynomials of finite even degree 2n, of the following form:

$$b_0 - b_1x^2 + b_2x^4 - \ldots + (-1)^n b_n x^{2n} = 0$$

with roots: $r_1, -r_1, r_2, r_2 \ldots, -r_n$.

He shows that the following holds:

$$b_1 = b_0\,(1/r_1^2 + 1/r_2^2 + \ldots + 1/r_n^2) \qquad (*)$$

All of this is quite regular mathematics. He then assumes that the same line of reasoning applies to polynomials of infinite degree. Thus the polynomial:

$$1 - x^2/3! + x^4/5! - x^6/7! + \ldots = 0,$$

with roots: $\pi, -\pi, 2\pi, -2\pi, 3\pi, -3\pi, \ldots$

As it is the series expansion of sin(x)/(x), will satisfy (*), thus

$$1/3! = 1/\pi^2 + 1/4\pi^2 + 1/9\pi^2 + \ldots,$$

or:

$$1 + 1/4 + 1/9 + \ldots = \pi^2/6 \qquad \text{QED (?)}$$

Although one might argue that arguments of this type could be called experimental, it is once again not at all clear in what sense. Both mathematical proofs and arguments are syntactic entities. Therefore they belong to the same category and this does not seem reconcilable with the idea of an experiment in the naive sense (and most likely in many extended senses).

Summarizing, it seems that it is very difficult to make any sense out of the idea of considering a computation (whether aiming for a numerical result or a visual image) as a form of mathematical experiment. At best one finds a partial analogy that breaks down at essential points. In all cases, what mathematicians apparently do is nothing but mathematics all of the time. They do not occasionally perform experiments. Having dealt with this category, let me turn to the other one.

3. Experiment = Real-World-Experiment

(a) I can remember reading years ago that the probability of two positive integers, chosen at random, being relatively prime is $6/\pi^2$. It seems that one R. Chartres, in about 1904, tested this mathematical result experimentally by having each of fifty students write down at random five pairs of positive integers. Out of the 250 pairs thus obtained, he found 154 pairs were relatively prime, giving a probability of 154/250. Calling this $6/x^2$, he found $x = 3.12$, while $\pi = 3.14159 \ldots$ (Honsberger, 1970, p. 3)

(b) Assertions like '2 + 3 = 3 + 2' or '101 is a prime number' appear to be directly validatable: the mathematician ascertains whether they correctly predict what he would experience by carrying out an experiment . . . (Rotman, 1988, pp. 13–14)

(c) In his book *The Mathematical Tourist* Ivars Peterson discusses the problem of Plateau: given a boundary curve B, what is the minimal surface S having B as its boundary? Mathematically this is a deep and difficult problem. Often, analytical methods are insufficient. There is however a simple way to find solutions, though not necessarily all solutions. Construct the boundary B in metal wire. Dip it in soapy water and a film will form having B as its boundary. Physics tell us that this film is a minimal surface. Hence, Peterson says: "They can explore shapes that are often too complicated to describe mathematically in a precise way. They can solve by experiment numerous mathematical problems associated with surfaces and contours." (Peterson, 1988, p. 48)

These examples clearly do not involve computers, but seem to be about 'real' experiments as real as one would naively associate with the notion of 'experiment.'

Consider the following example. I want to convince a rather stubborn pupil that 31 is a prime number. As he is not impressed with calculations and the like, I propose a game of matches. I ask him to arrange the matches in a rectangle. He tries to construct a rectangle with breadth 2. He does not succeed. He tries 3 and he does not succeed. As he continues, it gradually becomes clear that all his attempts fail. I now tell him that a prime number is a number such that the corresponding amount of matches do not allow you to construct a rectangle of whatever size (with breadth larger than one), and let us suppose that he agrees that 31 is indeed a prime number. The question is this: have I performed an experiment or not? Secondly, if it is an experiment, can it qualify as a *mathematical* experiment? The answer to the first question must be yes. The second question is the difficult one. Compare it with quote (c) above. Is it acceptable to call the experiment with the soap films a mathematical experiment? Should it not be seen rather as a *physical* experiment? In fact, is it not possible to consider any mathematical experiment, if such exists, as a physical experiment? I think the answer to the last question is no.

Consider any physical theory PT. Suppose that PT makes use of a mathematical theory MT. E.g., if PT is classical Newtonian mechanics, then a candidate for MT is the theory of vector spaces. In axiomatic format, PT will consist of a set of axioms some of which will be considered purely physical—say, Newton's third law—some of which will be considered purely mathematical—say, the associative law for vectors. If PT is applied to the real world, i.e., if we make a model for PT such that the elements in the model can be mapped onto 'real-world' elements, then, of course, the mathematical syntactical entities too will be mapped onto real-world entities. We have not turned the mathematical theory into a physical theory by mapping its syntactical elements onto physical objects. To put it differently, given a mathematical theory MT, among its models we will find purely abstract models—in fact, in some cases, e.g., as in set theory, the model itself is once again a mathematical structure—but we might also find physical models or real-world models. If this is indeed the case, then it is possible to perform a mathematical experiment where 'experiment' can be taken in the full sense of the word.

The skeptical reader might remark that although this leaves room for mathematical experiments, no serious implications should be drawn from them, as for the most part they are quite uninteresting. It is sufficient to look at the three examples above. (c) need not necessarily qualify as a mathematical experiment. Rather what one is testing is whether indeed nature obeys a principle of minimal energy. Suppose that a soap film turned out to have a completely unexpected geometrical form, then this would be devastating for the underlying physical theory but not necessarily for the underlying mathematics. (b) is rather trivial. It would be close to a miracle if an experiment of this type should produce a result that is unexpected. (a) finally is rather uninteresting. As we are dealing with probabilities, a large margin of error is acceptable as the example in fact shows. The result $\pi = 3.12$ does not force us to reconsider the value of π. Thus, although such experiments are perhaps possible, they are completely uninteresting and hence of no importance for mathematics as such.

The best way to answer this objection is by actually producing a counterexample, i.e., a case where a mathematical experiment turned out to play a crucial role. The subject is the tessellation of the plane (see Grünbaum & Shephard, 1987, for an excellent overview). The major problem is: can one list all possibilities? This general question reduces rather quickly to: what convex pentagons tile the plane? In 1918, Karl Reinhardt claimed but did not prove to have found a complete list. In 1968, Richard B. Kerschner stated that he had found a proof. It must be added that because of space limitations, he did not publish the proof itself. When Martin Gardner invited readers of the *Scientific American* to find counter-examples, someone found such a counter-example. This person was Marjorie Rice: ". . . an unlikely candidate for the role of mathematical innovator. A San Diego housewife and mother of five, she had no formal education in mathematics save a single course required for graduation from high school in 1939" (Rival, 1987, p. 41). Apparently, according to Rival, she was drawing pictures and putting them together mentally. In modern terminology, one would no doubt say that she was not doing mathematics. Clearly, it is quite acceptable to say she was experimenting (as understood here). Moreover, analysis of these additional tilings 'discovered' by Rice, produced further tilings, so it was a fruitful and productive experiment as well. Granted that one counter-example does not make a case, it does however show the possibility of *interesting* mathematical experiments.

Direct consequences of this interpretation are, on the one hand, that not all mathematical theories are subject to experiment. This seems quite obvious as the same holds for any scientific theory. Nothing prevents a physicist from constructing an elementary particle theory wherein the electron is twice as heavy as its measured value. Such a theory can have no consequences that are experimentally accessible. On the other hand, if mathematics does turn out to be experimentally accessible (in the full sense of the word), then this does not entail that mathematics is fallible. Although I tend to defend the idea of fallibility in mathematics, my reasons are different. If a mathematical experiment produces a result in conflict with the mathematical theory, it does signal a problem with the theory, but it does not signal that mathematics is wrong *as such*. I see no difference with the fact that no mathematician will abandon the idea that a good mathematical proof cannot be doubted, on the grounds that from time to time mathematicians produce proofs containing mistakes. The ideal notion of mathematical proof—in much the same way as the ideal notion of mathematical certainty remains largely unaffected by mistaken proofs. The same holds, I believe, for mathematical theories in conflict with experimental results. Phrasing it slightly differently, it is not because one accepts the possibility of mathematical experiments that one is forced to embrace a constructivist, empirical, Millian, or whatever philosophy of mathematics.

4. Conclusion

The major point of this paper is that at the present moment there is a lot of confusion concerning the notion of a mathematical experiment. The "experiment = computation" view, no doubt the most influential at the moment, is, as far as I can see, deeply misleading, whereas the "experiment = real-world-experiment" interpretation does seem to make sense, but here one might question whether it plays a truly fundamental role in mathematical research and whether it is a more important factor in mathematical practice today than in the past.

Finally, I must mention one last possible interpretation of a mathematical experiment. Suppose that there is such a thing as the Platonic universe of mathematical objects. And suppose that, as Gödel puts it: ". . . we do have something like a perception also of the objects of set theory, as is seen from the fact that the axioms force themselves upon us as being true. I don't see any reason why we should have less confidence in this kind of perception, i.e. in mathematical intuition, than in sense perception, . . ." (Gödel, p. 484). Granted that this form of "seeing" is possible, then, by extension, could one not perform an experiment in this world? Is all of this mere metaphor or are Platonists, strangely enough, the (sole) true experimenters in mathematics?

Notes

1. In order to save space, I have selected a few essential references.

2. This is, of course, a reference to Errett Bishop's paper "Schizophrenia in Contemporary Mathematics" (admittedly, Bishop refers to another type of schizophrenia).

3. See Tymoczko, 1986, for a discussion on the status of the (according to some, so-called) proof of the four-color theorem.

4. One must add here straightaway that this does not hold in general. There are some famous problems in mathematics, where all the 'evidence' pointed in one direction, yet, the proof, when finally found, pointed in exactly the opposite direction. The Mertens conjecture related to the Möbius function and the relation between $\pi(n)$—the number of primes less than n—and the logarithmic integral Li(n) as an approximation are two such examples. See Devlin, 1988, p. 210 and pp. 216–221. However, philosophers of science claim exactly the same thing.

5. Lack of space prevents me from saying more about this fascinating topic. See Brown, 1991, and Horowitz & Massey, 1991.

6. The suggestion that mathematical proofs can be classified according to such criteria as insight, quality, elegance, etc. has been taken up by Koetsier, 1991, esp. pp. 170–171 and myself, 1988. See also Fischer et al., 1993.

References

Appel, Kenneth, & Wolfgang Haken. "The Four-Colour Problem," in Lynn Arthur Steen (ed.), *Mathematics Today. Twelve Informal Essays,* Springer-Verlag, Heidelberg, 1978, pp. 153–180.

Bishop, Errett A. "Schizophrenia in Contemporary Mathematics," in Murray Rosenblatt (ed.), *Errett Bishop: Reflections on Him and His Research,* AMS, Providence, 1985, pp. 1–32.

Bochner, Salomon. *The Role of Mathematics in the Rise of Science.* Princeton University Press, Princeton, 1981.

Brown, James Robert. *The Laboratory of the Mind. Thought Experiments in the Natural Sciences.* Routledge, London, 1991.

Davis, Philip J., and Reuben Hersh. *The Mathematical Experience.* Birkhäuser, Boston, 1980.

Devlin, Keith. *Mathematics: The New Golden Age.* Penguin, Harmondsworth, 1988.

Dunham, William. *Journey Through Genius: The Great Theorems of Mathematics.* Wiley, New York, 1990.

Ernest, Paul. *The Philosophy of Mathematics Education.* Falmer Press, Basingstoke, 1991.

Fischer, Ronald, Sal Restivo, & Jean Paul van Bendegem (eds.). *Math Worlds: New Directions in the Social Studies and Philosophy of Mathematics.* State University New York Press, New York, 1993.

Gödel, Kurt. "What is Cantor's Continuum Problem," in: Paul Benacerraf and Hilary Putnam (eds.). *Philosophy of Mathematics. Selected Readings,* Cambridge University Press, Cambridge, 1983, pp. 470–485.

Grübaum, Branko, & G.C. Shephard. *Tilings and Patterns.* Freeman, San Francisco, 1987.

Honsberger, Ross. *Ingenuity in Mathematics,* New Mathematical Library, MAA, Washington, 1970.

Horowitz, Tamara, & Gerald J. Massey (eds.). *Thought Experiments in Science and Philosophy.* Rowman & Littlefield, Savage, 1991.

Kitcher, Philip. *The Nature of Mathematical Knowledge.* Oxford University Press, Oxford, 1983.

Koetsier, Teun. *Lakatos' Philosophy of Mathematics. A Historical Approach.* North-Holland, Amsterdam, 1991.

Lam, C.W.H. "How Reliable is a Computer-Based Proof?", *The Mathematical Intelligencer,* Vol. 12, 1, winter 1990, pp. 8–12.

Peterson, Ivars. *The Mathematical Tourist. Snapshots of Modern Mathematics.* Freeman, New York, 1988.

Rival, Ivan. "Picture Puzzling," *The Sciences,* Volume 27, number 1, 1987, pp. 41–46.

Rotman, Brian. "Toward a Semiotics of Mathematics," *Semiotica,* Vol. 72–1/2, 1988, pp. 1–35.

Steiner, Mark. *Mathematical Knowledge.* Cornell University Press, Ithaca and London, 1975.

Tymoczko, Thomas (ed.). *New Directions in the Philosophy of Mathematics.* Birkhäser, Boston, 1986.

Van Bendegem, Jean Paul. "Non-Formal Properties of Real Mathematical Proofs," in Arthur Fine and Jarrett Leplin (eds.), *PSA 1988. Volume One,* PSA, East Lansing, 1988, pp. 249–254.

Vivaldi, Franco. "An Experiment with Mathematics," *New Scientist,* 1688, Oct. 1989, pp. 46–49.