"Eureka" and Other Stories

Barry Mazur

June 29, 2005

Abstract

This is a text close to my oral presentation at the conference on *Mathematics and Narrative,* held in Mykonos, July 2005. For a fuller account, there is a *written text* that augments this; the written version is in two parts. *IT IS A STORY* is the first part, and is largely included here; it has no technical mathematics in it and is meant to be a discussion of the various possible roles of "story" in mathematics. The second part of the written text *IT IS A STORY (continuation)* gives details of the particular mathematical story (about ideas) that I only briefly allude to in my oral presentation. Much of this second part is material I drew from in a lecture I gave at the birthday conference in honor of Persi Diaconis, held in San Diego, in January of this year.

1 Theory versus Data

In an important, and on-going, direction of research, the *theory* governing the general shape of a certain aspect of number theory, and the extensive *computation* designed to get a more intimate view of the phenomena behind this, have—for a curious, but elementary, reason–difficulty being reconciled, one to another. This has led to some twists and turns in the expectations of mathematicians. This episode in contemporary mathematics brings with it two messages, one about the nature of *belief* in theoretical mathematics (theoretical heuristics if cogently supported, trumps data no matter how massive), and the other about *goals* in the framing of questions in number theory (our intellectual curiosity may not be satisfied even once we prove theorems that verify the asymptotic shape of important number theoretic data; we are, after all, always only confronted by finite—even if massive—data, that may need to be comprehended in a manner more explanatory than case by case computation).

I have been fascinated by this cliff-hanger, if I may call it that, for some time, and relish the opportunity of recounting it to you. To be sure, we are also mathematicians, and so we also enjoy any excuse to position things in their appropriate theoretical settings. Therefore, in the spirit of the conference, I will offer a general classification of stories, before sketching the particular one that has me in its grips, at present.

2 "Eureka"

I don't know whether or not Archimedes discovered his hydrostatic principle, the *principle of bouyancy,* while taking his bath. Nor do I know whether he immediately jumped from the tub shouting "Eureka," and ran home stark naked, dripping wet. But I do know that we all have our favorite stories that go along with accounts of mathematics; tales that help to explain, to dramatize, to teach, and even to shape in important ways, the mathematical material being recounted.

3 Linnaean classification

For a long time I have been intrigued, and also somewhat puzzled, by the motivating roles that stories (also poems) play in the unfolding of mathematical themes. So, for this conference I thought it might be fun to think a bit more about *story* per se, to attempt a kind of Linnaean classification of the various ends accomplished by narrative in scientific exposition, and finally to hint at a specific kind of mathematical story that grips me, at present.

4 Narrative as defined by a performative speech-act

Some years ago I read *Death in the Woods*, a short story of Sherwood Anderson[1]. It is a disarmingly simple account of a woman's death. In the woods. This piece of writing is like a pillow-feather, or a snowflake: it glints well in the sunlight, but I, at least, couldn't quite grasp it, the slightest breeze propelled it from my attention. A willow-the-wisp of a text. On the fourth reading I came across the declaration "It is a story" in the midst of the narrative. My eyes must have raced by those four words three other times, but I was happy to have finally registered them: I now had the author's collusion in being concerned with the question of whether or not it is, in fact, a story. "Yes, it is." he tells me.

What is a story? Is the previous paragraph a story?

5 Ends versus Means

Even before giving an answer to this in the context of stories in expository mathematics, we should be clear about whether the stories we will be considering are *ends* or *means.* In fiction, telling the story is the ultimate goal, and everything else is a means toward that goal. I suspect that even Sheherezade, despite her dangerous situation, and the immediate mortal purpose for her

[1]in preparation for attending a class given by the novelist Robert Boswell.

storytelling, would agree to this. It is also so in the Sherwood Anderson story, even where the story, one discovers, is about the author's relationship to the story.

In mathematical exposition any of its story elements are usually intended to serve the mathematical ideas: story is a means, the ideas are the end.

6 Kinds of storytelling that help explain the math

If, then, stories in mathematical exposition are a means, and not an end, *to what* are they a means: what do they accomplish? Let us try to throw together a provisional taxonomy of "kinds of storytelling" in mathematics, by the various possible answers to this question.

I feel that there are four standard forms, and also a fifth form: the one that I will be most specifically interested in, this hour. My names for the four standard ones are,

- *raisins in the pudding* (i.e., sheer ornaments),
- *origin*-stories (stories explaining some original motivation for studying the mathematics being described, this motivation being external to the development of mathematical ideas themselves),
- *purpose*-stories (again, other than those which describe a purpose within the context of mathematics itself),
- *drama, pure and simple* (either a personal drama, or a drama of ideas).

7 Raisins in the pudding

These are ornamental bits of story meant to provide anecdotal digressions or perhaps a certain amount of relief from the ardors of the main task of the exposition. At the least they are intended to add extra color. But the primary relationship of the stories or story-fragments in this category to the mathematical subject is ornament: they are not required to help in furthering–in any direct way–the reader's comprehension of the material, nor do they fit in as a part of the structure of the argument presented.

Archimedes' "Eureka" is such a raisin–at least when it isn't coupled to the more purpose-filled story of checking the gold content of King Hiero's crown–but other examples are plentiful, and any book in mathematics that has none of these ornaments risks being too aridly single-purposed to make for pleasant reading. The danger, though, is when the weight of ornament deflects the text from its primary purpose, as happens, in my opinion, in some recent popular books on math.

8 Purpose-stories

These, told in mathematical texts, are meant to answer the question: to what end–external to the development of mathematics itself– will this particular piece of mathematics be used? A practical end might be envisioned. Indeed, the great edifice of Applied Mathematics is devoted to mathematics strongly shaped by some particular issue or issues (e.g. in the "real world"). The term Applied Mathematics, it seems to me, covers a spectrum ranging from *commissioned problems* to *applicable methods*. But anywhere along this spectrum, Applied Mathematics has a purpose-story (whether explicit or implicit) at its very foundation.

Some purpose-stories do not connect with ends that could be labeled as practical. It often surprises me how even the tiniest sliver of such a purpose-story manages to focus the mind, and to clarify ideas. To begin to think about this, let us consider the "value-added" comprehension afforded to us, in Archimedes' *Sand-reckoner* (in which Archimedes establishes notation to describe very large numbers) when we read that the author wanted *to denote numbers larger than the number of grains of sand needed to fill up the universe.*

I believe that formulating this purpose– to denote numbers larger than the number of grains of sand in the universe– accomplishes three small, but important, pedagogical missions. First, it provides a very specific touchstone to judge whether the enterprise (finding notation for large numbers) will be deemed a success. A frame– admittedly a very large frame– has been put around the project, and we know what we are aiming for: according to Archimedes the number of grains tote up to 10^{63} (in our notation). Second, it is whispering something to us about the relationship between *matter* and *idea*. Here a language is being developed–Archimedes tells us–in the realm of idea, for quantities that are beyond those that could ever be realized by mere matter. Third, it emphasizes that this notational language can be precisely handled and understood, for the most specific of aims, even in the sublime range of numbers that have no material referent.

The above analysis of the pedagogical importance of the *sand* in Archimedes' treatise may be too simpleminded, but I feel that any bona fide purpose-story in a mathematical exposition will have some subtle effect on the presentation of the pure mathematical content, and what that effect is may be worth understanding.

9 Origin-stories

These usually answer the question of how the mathematician came to work on the material, how the ideas originated, They are critically different from purpose-stories because, at the very least, an origin-story leaves the endgame of the project open, although it claims to pinpoint the beginning-game. An example for this is given by the celebrated treatise of Archimedes (lost to us until 1906). This was the letter written to Eratosthenes that Archimedes called simply *The Method* and in which he proposed to offer what he called the "mechanical" method that preceded and led to many of his mathematical discoveries.

10 Drama, pure and simple

These comprise, of course, the mainstay of narrative. But in our mathematical context, the drama might involve personal crisis, as in these famous vignettes:

- the pythagorean perishing at sea (either because of human or divine intervention) because he revealed the secret that the ratio of the length of the diagonal of a square figure to the length of one of its sides is incommensurable,
- Young Evariste Galois rapidly writing his marvelous ideas in a letter to a friend the night before he died in a duel,
- Henri Poincaré realizing that there was a great lacuna to his work in celestial mechanics, as he was on a train coming back to Paris having accepted a prize for this very work. (In understanding this error Poincaré discovered the phenomenon of *homoclinic points.*)

or it might involve more purely a crisis of ideas or it might be a complex combination of both, as exemplified by the story of the life and thought of Kurt Gödel.

The phrase "crisis of ideas" has crept into the previous paragraph. What in the world can such a thing be?

11 An amateur taxonomy of intellectual crises in mathematics

When some established overarching framework, or vocabulary, or procedure of thought, is perceived as *inadequate in an essential way,* or as *not meaning what we think it means,* then Kuhn-like trouble is brewing. Here are two basic ways in which such trouble comes to light.

12 Problems related to the existence of concepts

The single oldest theme that offers "intellectual crises in mathematics" is *ontology.* By this, I mean when we are forced to contemplate the possibility of existence of a mathematical entity, or concept, that defies the heretofore accepted criteria for existence. For example,

- The ratio of the length of the diagonal of a square figure to the length of one of its sides is not "the ratio of a number to a number." Pappus, in his *Commentary on Book X of Euclid's Elements,* says about this pythagorean discovery,

  > [The pythagoreans] sought to express their conviction that firstly, it is better to conceal (or veil) every surd, or irrational, or inconceivable in the universe, and, secondly, that the soul which by error or heedlessness discovers or reveals anything of this nature which is in it or in this world, wanders [thereafter] hither and thither on the sea of non-identity (i.e., lacking all similarity of quality or accident), immersed in the stream of the coming-to-be and the passing-away, where there is no standard of measurement.

- The doctrine of the *indivisible line* is attributed to Xenocrates of Chalcedon. This concept, indivisible line, i.e., an element of a line segment *so small that it has no parts* is reminiscent of the modern *infinitesimal* or *tangent vector.* Here is a shred of the ancient argument *for* its existence:

  > ... if there is an Idea of a line, and of the Idea is *first* of the things called by its name:—then, since the parts are by nature prior to their whole, the Ideal, Line must be indivisible.

  The standard ancient argument *against* the existence of indivisible lines, by the way, is to quote the proposition of Euclid that any line segment can be bisected; and how in the world can you bisect an indivisible line?

- Whether or not cubic radical expressions such as

  $$\sqrt[3]{5+2\sqrt{-1}}+\sqrt[3]{5-2\sqrt{-1}}$$

  exist preoccupied Rafael Bombelli for the twenty year period during which he wrote his wonderful *L'Algebra.*

13 Problems related to the "limits of reason"

This theme has only come into prominence in discussions of mathematics since Kant. But the issue forced itself front and center into the language of mathematics itself, in a critical way, at the time of the Frege-Russell correspondence. The encounter between Frege and Russell provides as rich fodder for mathematical narrative as anything I know, but it is too well-known a story for me to revisit it here. To be sure, it, together with the goading of L.E.J. Brouwer, provoked Hilbert to offer a reformulation of the entire edifice, and with that, perhaps, the mission, of mathematics.

This is laid out in breathtaking audacity in Hilbert's foundational articles and essays, notably *On the Infinite,* where Hilbert thinks that he is coming to the rescue of Cantor's theory. Here is what he writes.

> From the most diverse quarters extremely vehement attacks were directed against Cantor's theory itself. The reaction was so violent that the commonest and most fruitful notions and the very simplest and most important modes of inference in mathematics were threatened and their use was to be prohibited. [...] Let us admit that the situation in which we presently find ourselves with respect to the paradoxes is in the long run intolerable. Just think: in mathematics, this paragon of reliability and truth, the very notions and inferences, as everyone learns, teaches, and uses them, lead to absurdities. And where else would reliability and truth be found if even mathematical thinking fails?

The error that mathematics fell into, as Hilbert analyzes it, is as follows.

> ... we accepted arbitrary abstract notions, in particular those under which infinitely many objects are subsumed.

Hilbert conceded (in effect, to Brouwer) that "we," (i.e., the mathematical community) were using logical inference in an illegitimate way, and we did not respect *necessary conditions* for its use.

> And in recognizing that such conditions exist and must be respected we find ourselves in agreement with the philosophers, especially with Kant. Kant already taught–and indeed it is part and parcel of his doctrine–that mathematics has as its disposal a content secured independently of all logic and hence can never be provided with a foundation by means of logic alone; that is why the efforts of Frege and Dedekind were bound to fail.

The denouement of this meditation of Hilbert is well known. To defend mathematics, ("No one shall be able to drive us from the paradise that Cantor created for us.") Hilbert goes on to conceive the notion of *formal system* as a mathematical entity in its own right, thereby being able to give a precise (albeit a relative) sense to the concepts of *consistency, provability, undecidability.* The implicit goal of the new project that this gives birth to is to establish, once and for all, a universally usable, self-consistent, formal system which is a fount for all demonstrations in mathematics. That such a precisely formulated a goal can be as wildly successful, and at the same time as wildly unsuccessful, as this one turned out to be, makes for quite a story.

It is wildly successful because Hilbert's format of formal systems, axiomatic systems (coming with the vocabulary of *sets* and *mappings of sets*) has become the lingua franca of modern mathematics. This format has allowed for the great advances mathematics has enjoyed in modern times. Moreover, we teach this lingua franca so well to our students that many of them would be puzzled to think that mathematics might have been conceived of, in earlier times, somewhat differently.

It is, of course, also wildly unsuccessful—given its Hilbertian mission—as Gödel showed.

14 Before and after Hilbert

One major side-effect of the Hilbert program is that it seriously muted—or at least transmuted—the nature of *ontological* drama in mathematics. If you read the older mathematical literature you sometimes encounter explicit expression of agonies of doubt regarding the existence of a certain concept, as when Cardano claimed that it would be foolish to consider fourth powers, or higher powers, because "nature doesn't permit it." One also encounters in the older literature flurries of optimism about the existence of entities, such as when—as alluded to above—Bombelli comes (very tentatively) to the conclusion that his strange cubic radicals *exist.*

In the literature after Hilbert, the tenor of these kinds of reflections has changed. We all know how to package mathematical thought into axiomatic systems. The type of questions that arise, that stand in for ontological concerns, either have their expression in the technical vocabulary of formal systems: Is it consistent? Does it admit a model? Or else, the questions focus on the explanatory power of the axiomatic system: Is it germane? Is it leading us in fruitful directions?

15 Narratives of mathematical ideas unfolding

The fifth category of story in mathematical exposition, the one far more difficult, in fact, to categorize than the others, is the kind of narrative that is traced out by the ideas themselves as they unfold. The fundamental propelling force of any story is made palpable by the energy behind the question "and then?" that the listener asks. The yearning to do this–to ask "What happens next?"–in the realm of ideas, the impulse usually called *intellectual curiosity*, is itself a most curious thing. To begin, it presumes that there is a "next." Narrative depends critically on our being creatures in time. Any argumentation requires this, as well. Given X we proceed to Y, and thereby weave an ordered succession through an otherwise atemporal intellectual realm. What this entails deserves to be thought through, but I won't try to do that here[2]. Suffice to say that a pattern of suspense, however muted, must be felt; and the standard engine that drives suspense–namely, some form of conflict or tension–must be present.

[2] Jean-Francois Burnol recently sent me an e-mail saying that in doing this we are like voyagers exploring a spider web who will "par la force des choses" create a temporal ordering which can be but a pale incarnation of the full structure.

16 The phenomenologist's dilemma

The essential tension in the mathematical story that interests me—I have time only to allude to it, rather than go into it, to in my talk—is between two basic elements that are usually present in any live piece of mathematics. On the one hand, we have a network of heuristics and conjectures. On the other hand we have massive data accumulated to exhibit instances of the phenomena. Generally, we would expect that our data supports our conjectures; and if not, we lose faith in our conjectures. But here there is a somewhat more surprising interrelation between data and conjecture: they are not exactly in open conflict one with the other. But it may very well be that until we actually *demonstrate* our conjectures, no data that we will accumulate, however massive it may appear, will give even lukewarm comfort to the conjecturers. As a result, our expectations about this chapter of mathematics have had something of a see-saw history. This conflict (or pseudo-conflict) raises the question of whether we as mathematicians may, at times, face a situation where the substance we study has one shape *asymptotically*, and yet all computational evidence elucidating this substance, even up to the very large numbers that computers today, or in our lifetime, can compute, seem consistent with the possibility that the data have a different asymptotic shape.

Here, then, is the little mathematical story I am currently gripped by, where data– massive though it may be– does not give indubitable confirmation of the heuristics that a large number of mathematicians actually believe; and, nevertheless, the heuristics are indeed believed.

17 Counting rational points on elliptic curves

The topic is rational points on elliptic curves.

Elliptic curves have been connected to many of the recent advances in the subject, Fermat's Last Theorem being among them. But elliptic curves have played a major role in mathematics, from early on: in mechanics, abelian integrals, Riemann surfaces, the theory of doubly periodic analytic functions, and automorphic forms. These elliptic curves amply repay the obsessive interest that mathematicians have for them: the tightness of their structure gives us unexpected leverage in applying them to solve a host of problems. Elliptic curves seem to be designed to teach us things, and nowhere less than in arithmetic.

Rational points are the gems of the arithmetic theory. How "many" rational points are there? *Elliptic curves* always have one rational point. But they have the peculiar feature (among other classes of curves) that it is possible for them to have only finitely many rational points, or infinitely many, and on this dichotomy hangs much deep arithmetic. So, it is natural to ask for percentages.

18 What is the probability that a given elliptic curve has infinitely many rational points?

To get the story here, take a look at the second part of my written text for this conference. But the stark question is a curious one:

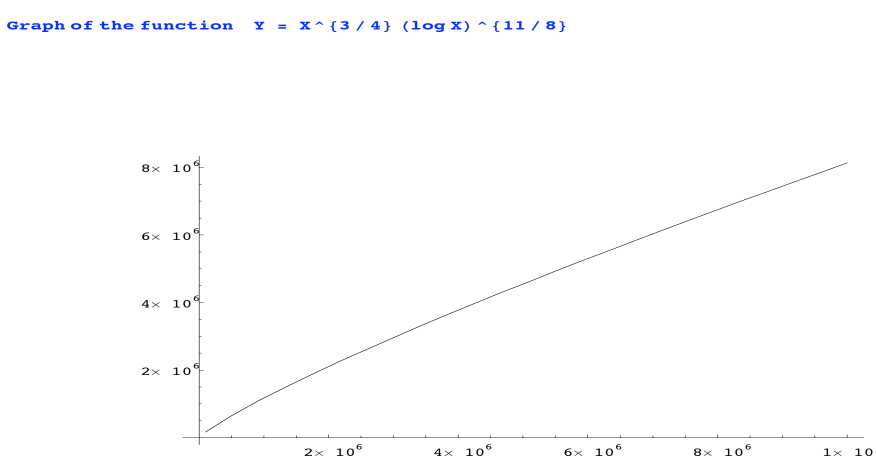
- Is the probability about 2/3?
- Is the probability 1/2?

If you believe the data—and there has been truly massive data accumulated here, for elliptic curves of conductor up to 10^8—you would say that it is roughly 2/3 for the numerical evidence unwaveringlu supports this percentage, up to the extensive range that has been calculated.

But no one, currently, believes this to be the true percentage! For the theoretical evidence offers up 1/2 as the correct probability. There has been, though, in the pas history of this subject considerable vacillation of opinion.

19 graphs that look very linear

With minimal technical langauge, here is the crux, the simple mathematical issue that is responsible for the sense of conflict here. Consider this graph in the range of numbers X from $X = 10^6$ to $X = 10^8$. If this graph came to you as some kind of statistical raw data, might you not form the opinion that the curve that fits that data is a straight line?



That the function $Y = X^{3/4} \log(x)^{11/8}$ visibly and significantly different asymptotically from any linear function of X—as any first-year calculus student knows—can have graphs with surprisingly similar "looks," even at relatively high values of X, accounts for this persistent discrepancy between the two probabilities in the the mathematical question we have posed in the previous section. It is also responsible for an extraordinary amount of vacillation of opinion among the experts about the asymptotic behavior of this mathematical event. It also makes it very unlikely–until computers get significantly more powerful than they are today–that we will ever see data related to the problem described by that graph, that convincingly confirms our theoretical expectation. Of course, we might eventually *prove* our expectation.

20 Is it a story?

Oh yes, it is. It underscores dramatically—as did the *incommensurables* first to the pythagoreans, the *indivisible line* to Xenocrates, and *Bombelli's cubic radicals* to Bombelli—how wonderful, how surprising, and how strange it continues to be: the interplay of *data* and *idea* in mathematics.