CHAPTER 7

Vividness in Mathematics and Narrative

TIMOTHY GOWERS

Is there any interesting connection between mathematics and narrative? The answer is not obviously yes, and until one thinks about the question for a while, one might even be tempted to say that it is obviously no, since the two activities seem so different. But on further reflection, one starts to see that there are some points of contact. For example, to write out the proof of a complicated theorem one must take several interrelated ideas and present them in a linear fashion. The same could be said of writing a novel. If the novel is describing a series of events, then those events will have a natural order, but there is much more to a good novel than "A happened, then B happened, then C happened, then . . . .," and the order in which information is revealed to the reader is often not chronological. Likewise, in a mathematical presentation statements come in a logical order (which may be partial rather than total), and this order frequently differs both from the order in which the statements are discovered and from the order in which they should be presented if they are to be understood most easily. The result of all this is that mathematicians who wish to communicate their ideas effectively face many of the same challenges as novelists.

This essay is not a general discussion of connections between mathematics and narrative. Rather, I want to focus on one particular quality, which I shall call "vividness," that passages of narrative can have to a greater or lesser extent. I shall argue that mathematical explanations can have or lack this quality as well, and that the causes are similar. To give some idea of what I am talking about, here is a description of the beginning of an academic year.

> It is September again, and the campus, which has been very quiet for the last couple of months, is suddenly full of cars bringing students back after their vacation. The parents are well-to-do; the cars are expensive and packed with things that the students will need during the term, though not all these items are strictly necessary. There is a general, if unspoken, sense among the parents that they are all from the same sector of society, with similar attitudes and similar ways of life. This makes them feel comfortable, and perhaps even a little smug. At first their sons and daughters are a bit shy, which causes some of them to be

quiet and others to be overexcited. But this will wear off very soon: the transition will be forgotten and the term will have properly begun. The weather is typical for the time of year, with just a hint of the change that will take place over the next three months.

And here is another one:

> The station wagons arrived at noon, a long shining line that coursed through the west campus. In single file they eased around the orange I-beam sculpture and moved toward the dormitories. The roofs of the station wagons were loaded down with carefully secured suitcases full of light and heavy clothing; with boxes of blankets, boots and shoes, stationery and books, sheets, pillows, quilts; with rolled-up rugs and sleeping bags; with bicycles, skis, rucksacks, English and Western saddles, inflated rafts. As cars slowed to a crawl and stopped, students sprang out and raced to the rear doors to begin removing the objects inside; the stereo sets, radios, personal computers; small refrigerators and table ranges; the cartons of phonograph records and cassettes; the hairdryers and styling irons; the tennis rackets, soccer balls, hockey and lacrosse sticks, bows and arrows; the controlled substances, the birth control pills and devices; the junk food still in shopping bags—onion-and-garlic chips, nacho thins, peanut creme patties, Waffelos and Ka-booms, fruit chews and toffee popcorn; the Dum-Dum pops, the Mystic mints.
>
> I've witnessed this spectacle every September for twenty-one years. It is a brilliant event, invariably. The students greet each other with comic cries and gestures of sodden collapse. Their summer has been bloated with criminal pleasures, as always. The parents stand sun-dazed near their automobiles, seeing images of themselves in every direction. The conscientious suntans. The well-made faces and wry looks. They feel a sense of renewal, of communal recognition. The women crisp and alert, in diet trim, knowing people's names. Their husbands content to measure out the time, distant but ungrudging, accomplished in parenthood, something about them suggesting massive insurance coverage. This assembly of station wagons, as much as anything they might do in the course of the year, more than formal liturgies or laws, tells the parents they are a collection of the like-minded and the spiritually akin, a people, a nation.

Obviously the second account is incomparably more vivid than the first, and better for many other reasons too. It is the virtuoso opening of Don DeLillo's cult novel *White Noise*. The first passage was written by me. (In my defense, its blandness was intentional.) Later, I shall discuss an important respect in which the two passages differ, which plays a large part in our perception of the vividness of the second.

But before I do that, let me give two presentations of the mathematical notion of a group.

1. What Is a Group?

1.1. First Answer

Let X be a set. A *binary operation* on X is just a function ϕ from $X \times X$ to X. It is customary to use a symbol such as $\circ$ for this function and to write $x \circ y$ instead of $\circ(x, y)$. A binary operation $\circ$ on a set X is said to be *associative* if $x \circ (y \circ z) = (x \circ y) \circ z$ for any three elements x, y, z of X. An *identity element* for $\circ$ is an element e of X such that $e \circ x = x \circ e = x$ for every x in X. Note that an identity element is unique if it exists: if e and f are identity elements, then $e = e \circ f = f$. If e is an identity element and x belongs to X, then y is said to be an *inverse* for x if $x \circ y = y \circ x = e$. Inverses are again unique if they exist (assuming associativity): if $y \circ x = x \circ y = z \circ x = x \circ z = e$, then $z = z \circ e = z \circ (x \circ y) = (z \circ x) \circ y = e \circ y = y$.

A *group* is a set X together with a binary operation $\circ$ such that the following axioms are satisfied.

1. $\circ$ is associative.
2. There is an identity for $\circ$.
3. Every element of X has an inverse.

Most of the time in group theory we write xy instead of $x \circ y$: it is to be understood that this is a useful shorthand for $x \circ y$, where $\circ$ is the binary operation used to define the group. (This is often referred to as the *group operation*.)

1.2. Second Answer

From early childhood, we are all familiar with the idea of symmetry. On looking in a mirror, we note with amusement that when we move our right arm, our reflection appears to move its left arm—an effect that depends on the fact that human bodies look approximately the same if they are reflected in a vertical plane that separates them down the middle. (The "paradox" that we see left-right inversion but not up-down inversion in a mirror is explained by the fact that the human body is not approximately the same if we reflect it in a horizontal plane that goes through the waist.)

The mathematician's take on symmetry is slightly different. To a mathematician, symmetry is not so much a static *property* of an object but rather something you can *do* to an object. For example, take an equilateral triangle. The layperson might say that it is quite a symmetrical shape: it is symmetrical about three lines of reflection, and has rotational symmetry as well. A mathematician would say that the equilateral triangle *has symmetries* rather than *is symmetrical.* These symmetries are the three possible reflections, the two possible rotations (clockwise through 120 degrees and anticlockwise through 120 degrees), and the seemingly pointless "identity transformation," which consists in doing nothing at all. Thus, to a mathematician, a symmetry of a shape means something you can do to that shape that leaves it looking the same afterward as it did before. (Imagine, for instance, that on your kitchen table there is a plastic equilateral triangle. You leave the room for a while, and unbeknownst to you a friend rotates it through 120 degrees about its center. When you come back, you won't notice any difference.)

A simple observation that turns out to have ramifications throughout all of modern mathematics is that if you do two symmetries, one after another, the result is a third symmetry. For example, if you reflect an equilateral triangle and then rotate it, the result turns out to be the same as if you had reflected the triangle about a different line. We call the new symmetry the *composition* of the other two symmetries. Note that if you rotate a triangle through 120 degrees clockwise and then through 120 degrees anticlockwise, you will end up having done nothing at all. This is why the identity transformation is important: without it, we could not say that the composition of two symmetries was a symmetry.

Now let us look at another example of composition. Consider the following six functions: $f_0(x) = x$, $f_1(x) = 1/(1-x)$, $f_2(x) = 1 - 1/x$, $g_1(x) = 1/x$, $g_2(x) = 1 - x$, $g_3(x) = x/(x-1)$. If you compose any two of these functions, you will get a third. For instance, $f_1(g_1(x)) = 1/(1 - 1/x) = x/(x-1) = g_3(x)$. More interestingly, if you play around with these functions and also with the symmetries of an equilateral triangle, you start to realize that there are close similarities between the two. For example, if we compose anything with f_0, we don't change it. So f_0 is very like the identity transformation of the equilateral triangle. Also, if we compose f_1 with itself, we get $f_1(f_1(x)) = 1/(1 - 1/(1-x)) = (1-x)/(-x) = 1 - 1/x = f_2(x)$. If we compose this with f_1 again, we get $f_1(f_2(x)) = 1/(1 - (1 - 1/x)) = 1/(1/x) = x = f_0(x)$. That is, doing f_1 three times gets you back to where you started, so f_1 is a bit like a rotation through 120 degrees.

We can pursue this line of thought. It is easy to see that if you do g_1 or g_2 twice, then you get back to where you started, and a small calculation shows that the same is true of g_3, which suggests that they might be like reflections.

There appears to be an analogy between these two situations, but what exactly is that analogy? We can get a precise answer to that question if we draw up a "multiplication table" for the symmetries of an equilateral triangle. The letter e in table 7.1 stands for the identity transformation, ρ_1 stands for a rotation through 120 degrees anticlockwise, and ρ_2 stands for a rotation through 120 degrees clockwise. The letters σ_1, σ_2, and σ_3 stand for the three possible reflections, notated in such a way that the axis of σ_2 is obtained from the axis of σ_1 by a 120 degree rotation anticlockwise. The entry in the row marked ρ_2 and the column marked σ_1 is σ_2. That is because $\rho_2 \circ \sigma_1 = \sigma_2$: to put that another way, if you first do σ_1 and then do ρ_2, the result is the same as doing σ_2. (The rules for composition dictate, slightly confusingly, that the transformation that appears on the right is the one that you do first.)

Table 7.1. The symmetries of an equilateral triangle.

$\circ$	e	ρ_1	ρ_2	σ_1	σ_2	σ_3
e	e	ρ_1	ρ_2	σ_1	σ_2	σ_3
ρ_1	ρ_1	ρ_2	e	σ_3	σ_1	σ_2
ρ_2	ρ_2	e	ρ_1	σ_2	σ_3	σ_1
σ_1	σ_1	σ_2	σ_3	e	ρ_1	ρ_2
σ_2	σ_2	σ_3	σ_1	ρ_2	e	ρ_1
σ_3	σ_3	σ_1	σ_2	ρ_1	ρ_2	e

With the help of table 7.1, one can say exactly what the analogy is. Suppose we were to draw up a multiplication table for the functions f_0, f_1, f_2, g_1, and g_2 discussed earlier. Then we would obtain exactly the same table.

Of course, this isn't quite true, as it would be full of f s and g s. But it would become true if we gave the functions different names: if we rename f_0 as e, f_1 and f_2 as ρ_1 and ρ_2, and g_1, g_2, and g_3 as σ_1, σ_2, and σ_3, then it really is the case that the above table is the "multiplication table" for the six functions.

Thus, what the two situations have in common is the structure of the multiplication table. If we concentrate just on the multiplication table and forget about the nature of the objects that make it up, then we arrive at the abstract notion of a group.

It is worth pointing out that it is not a complete coincidence that the two multiplication tables are the same. One can think of the three numbers 0, 1, and ∞ as the vertices of a kind of triangle. The function $f_1(x) = 1/(1 - x)$ sends 0 to 1, 1 to ∞, and ∞ to 0, so from this point of view it really is behaving like a rotation that gets you back to where you started if you do it three times. Similarly, the function $g_1(x) = 1/x$ leaves 1 where it is and swaps 0 with ∞, so it is like a reflection. (With the help of Möbius geometry, one can even think of these functions as genuine rotations and reflections.)

These two introductions to the notion of a group are completely different in style, and each has its advantages. The first is clear, concise, and unambiguous. In a sense, it tells you everything you want to know. But it is also flat and mechanical, and it gives one absolutely no reason to be interested in groups. The second is much longer and makes an attempt to show how the notion of a group arises naturally from more basic notions such as symmetry and the composition of functions. However, after two pages it still has not said what a group is. (Of course, this last problem is easily remedied: the only reason I did not remedy it was that I did not want to be too repetitive. If I were to continue the account, I would observe that function composition was associative, and that in both examples we had an identity and inverses, and I would then say that having those properties was what we meant when we said that the symmetries or functions formed a group.)

A significant advantage of the second introduction is its vividness. It allows one to "see" groups in a way that the much more formal first account does not. But how is this vividness achieved? The answer is very simple: no abstract concept is introduced until an example that illustrates it has already been discussed. For example, the abstract notion of a symmetry, as a transformation that leaves a shape unchanged, is introduced only after extensive discussion of mirror images, rotations, and so on. Composition is not defined until a composition of symmetries has been discussed. And the abstract notion of a group is not defined until two different but isomorphic groups have been discussed in considerable detail.

If we look back at the passages with which this article began, we can see a similar but not quite identical phenomenon. Contrast the following two excerpts, one from the first passage and one from the second.

> There is a general, if unspoken, sense among the parents that they are all from the same sector of society, with similar attitudes and similar ways of life. This makes them feel comfortable, and perhaps even a little smug.
>
> The parents stand sun-dazed near their automobiles, seeing images of themselves

in every direction. The conscientious suntans. The well-made faces and wry looks. They feel a sense of renewal, of communal recognition. The women crisp and alert, in diet trim, knowing people's names. Their husbands content to measure out the time, distant but ungrudging, accomplished in parenthood, something about them suggesting massive insurance coverage. This assembly of station wagons, as much as anything they might do in the course of the year, more than formal liturgies or laws, tells the parents they are a collection of the like-minded and the spiritually akin, a people, a nation.

The first passage wastes no time: it just lays out the abstract idea. The second passage waits until the last few words before explicitly expressing the same idea, but by that time the idea has been conveyed by means of numerous small concrete details: the suntans, the bearing of the women, the fact that they know names, the "distant but ungrudging" husbands, the insurance coverage, the station wagons. And the cumulative effect of all this is that one can "see" these parents in a way that one cannot see the parents in the first passage.

Why should this be? I am not an expert in literary criticism (still less the more specific school of criticism known as reader response theory) or cognitive science. However, something like the following account seems obviously correct. Our brains contain a mass of information that is linked by means of a vast web of associations. Most of these associations are built up as a result of years of untidy, concrete, specific human experience rather than tidy, abstract, general reasoning. Therefore, prose that concentrates on concrete and specific details is usually much better at triggering associations than prose that is abstract and general. In the hands of a skilled novelist, this effect can be remarkably powerful. By the time DeLillo tells us that the parents are a collection of the like-minded and spiritually akin, we have them in our heads, his clever choices of words having caused our brains to retrieve memories of experiences we have had of seeing similar people (or perhaps complicated amalgams of such experiences), which tricks us into feeling as though we are actually there witnessing the scene he describes. When you read the passage I wrote, you have to create the parents for yourself. If you do not have the energy or imagination to do a good job, then they will exist in your mind in a vague, abstract, and above all unvivid form.

The analogy with mathematical presentations is almost too obvious to be worth spelling out. When we read a mathematical text, we come to it with our brains already full of a mass of associations built up as a result of years of untidy, concrete, specific mathematical experience. Therefore, a mathematical presentation that includes plenty of concrete examples is usually much better at triggering associations than a mathematical presentation that is purely abstract. By the time the second treatment of groups actually gets to the point where groups are about to be defined, we have some groups in our heads already. In a sense, all that is needed at this point is to say, groups are things like *that*. When you read the first treatment, you have to create some examples for yourself. If you can't face it, then you won't really understand what a group is.

A couple of years ago I wrote a blog post in which I argued, for roughly the reason just given, that when explaining a mathematical concept it was a good idea not just to give examples but to give them *first*. Usually, once you know that the abstract concept *is* an abstraction of certain examples, the concept becomes easy to remember, at least if the examples are well chosen. For instance, if you know that the rational numbers form a field and so do the integers modulo a prime p, then you can think of a field as the algebraic structure given by all the axioms that those two structures satisfy rather than as the algebraic structure given by a list of axioms that is long and apparently arbitrary. (As a mathematical definition that is unsatisfactory, but as a mnemonic it is pretty good.) I was surprised to discover quite a lot of opposition to the idea. Several people commented that if they are presented with examples without knowing what the examples are illustrating, they feel as though their feet are not on solid ground. Better, in their view, was to present the abstract definition and then to follow it immediately with examples.

It may be that what works best is different in different contexts. For instance, perhaps in a less formal treatment of a subject it is better to present examples first, and in a lecture course it is better to present examples immediately after an abstract definition (which is probably the approach that most lecturers take). I myself have exactly the opposite reaction to that I have just described: I feel as though my feet are not on solid ground if I am presented with an abstract idea without having some examples to relate it to. But this debate, interesting though it is, does not affect the point I want to make here, which is that a concrete, examples-driven mathematical discussion will be more vivid than a purely abstract one. Similarly, although I personally prefer narrative that (most of the time at least) follows the well-known instruction, "Show, don't tell," I recognize that telling has its place. But that does not alter the fact that showing is nearly always more vivid.

Let me give two more examples. Suppose I wanted to teach the rule $x(y + z) = xy + xz$ to a class of twelve-year-olds. I could just say, "One of the most important rules in algebra is that $x(y + z) = xy + xz$," but then the only people who understood me would probably be those who were already consciously aware of the rule. Others might not realize that $a(b + c) = ab + ac$ and would probably have considerable difficulty understanding that $3(x - 2y) = 3x - 6y$ or that $(x + y)(z + w) = x(z + w) + y(z + w)$. But if instead I started by asking why it is obvious that $30 + 60 = 90$, and discussed in detail how to work out 23×36, then the distributive law and its generalizations could be thought of as "The kinds of things I was doing in those calculations." Somehow, the rule comes to life.

The second example comes from fiction, in the extreme sense that the piece of fiction I am referring to is itself fictional. Suppose that a novelist wanted to convey the reaction of a man who has just heard that his son has been killed. The abstract approach would be to say something like, "He was filled with an intense grief." The more concrete approach would be to describe the man from the outside, so to speak. For instance, "He turned and looked out of the window, where his next-door neighbor was mowing the lawn. We spent the next ten minutes in silence." The concrete approach is more vivid here because it is very easy for the brain to conjure up a window (perhaps from some very early moment in childhood when we first learned the word *window*—in my case, when I examine the mental image I have of the house, I realize it is a house that I lived in between the ages of about two and five) and the view out of it. We also feel that we are watching the man, wondering what he is thinking, contrasting what must be going on in his head with the innocent contentment, or so we suppose, of the next-door neighbor, and so on. It is much less easy to conjure up intense grief out of thin air, though of course it will be easier for those who have had the misfortune to feel intense grief themselves.

One difference that might seem important between the way that information is conveyed in mathematics and the way it is conveyed in literature is that even if examples are very helpful in mathematics, they cannot *replace* abstract discussion in the way that they can in literature. "Show, don't tell" would be a very strange piece of advice to a mathematician: the strongest injunction of that kind that is reasonable is "Show first, then tell."

But even this distinction is a matter of degree. Recall the end of the passage from *White Noise*: "This . . . tells the parents they are a collection of the like-minded and the spiritually akin, a people, a nation." Having beautifully shown us his point, Don DeLillo tells it to us. Conversely, there are circumstances where an abstract discussion in mathematics adds nothing of value once an example has been discussed. This takes a little more effort to illustrate.

2. How to Calculate Highest Common Factors

2.1. Demonstration by Example

What is the highest common factor of 247 and 403? To answer this question, we could just work out all the factors of 247 and all the factors of 403, and pick the largest number that is a factor of both. But there is a much more efficient method due to Euclid. Note that any factor of 247 and 403 must be a factor of their difference $403 - 247$, which equals 156. Conversely, any factor of 247 and 156 must be a factor of 247 + 156, which equals 403. Therefore, the highest common factor of 247 and 403 is the same as the highest common factor of 247 and 156.

We can repeat this observation. Exactly the same reasoning shows that the highest common factor of 247 and 156 is the same as the highest common factor of 156 and 247 − 156, which equals 91. Then

$$\mathrm{hcf}(156, 91) = \mathrm{hcf}(91, 65) = \mathrm{hcf}(65, 26) = \mathrm{hcf}(39, 26)$$
$$= \mathrm{hcf}(26, 13) = \mathrm{hcf}(13, 13) = 13.$$

At each stage of the process above, we replaced the larger number by the difference of the two numbers. But there is a further idea that can save a great deal of time. Suppose that we reached the two numbers 137 and 2511. If we followed the process above, then we would find ourselves subtracting 137 several times, but we do not have to. Instead, we can work out the largest multiple of 137 that is less than 2511 and subtract that. A small calculation reveals that this multiple is 17 × 137 = 2429. Then any factor of both 137 and 2511 will be a factor of 2511 − 17 × 137, which is equal to 82, and conversely any factor of both 137 and 82 is a factor of 137 + 17 × 137, which equals 2511. Therefore, in a single (slightly more complicated) step we can replace the pair (2511, 137) by the pair (137, 82). Using the same idea in the previous calculation, we would have gone straight from the pair (65, 26) to the pair (26, 13), the number 13 coming from subtracting 2 × 26 = 52 from 65.

It is an unfortunate fact of mathematical life that algorithms are often hard to describe with complete precision, at least if you want them to be understood. Here is how one might go about it in the case of the Euclidean algorithm (which is less hard than most).

2.2. General Description

Let x and y be two positive integers, and suppose that $x \geq y$. Then we can write $x = qy + r$ for some positive integer q and some integer r with $0 \leq r < y$. Then any factor of y and r is a factor of x, and any factor of x and y is a factor of r (since $r = x - qy$), from which it follows that $\mathrm{hcf}(x, y) = \mathrm{hcf}(y, r)$. Since $r < y$, if we iterate this process it must terminate, and it can do so only if it reaches a pair of the form $(a, 0)$ for some positive integer a. Since $\mathrm{hcf}(a, 0) = a$, we then know that $\mathrm{hcf}(x, y) = a$.

I would not want to say that a description of this second kind is not desirable. Indeed, a precise description like that is very useful if one wishes to generalize the Euclidean algorithm to other situations, such as polynomials, and essential if one wishes to consider in detail the fascinating and important ways in which it can *fail* to hold (in more general rings than the integers). However, what I do maintain is that it is possible to teach somebody how to apply the Euclidean algorithm by showing them a couple of examples and not bothering to give them the general description. And that is all I need to establish in order to make the point that it is sometimes possible, even in mathematics, to show without telling.

◆ ◆ ◆

I said this would not be a general discussion of connections between mathematics and narrative, but I cannot resist a brief discussion of one or two other literary devices. For instance, is there any place in mathematics for metaphor, which is clearly of central importance in literature? Simile can certainly be useful. For example, it is helpful to think of a module as "like a vector space." And writers of popular mathematics books often try to find apt metaphors or similes to explain complicated mathematical concepts to the layperson. But again we have the phenomenon that whereas in literature the figure of speech is usually sufficient on its own, in mathematics it usually isn't. To make sense of the comparison between modules and vector spaces, we need to say more: a module is like a vector space, but with scalars that belong to a ring rather than to a field. Perhaps with an extra qualification we could upgrade the simile to a metaphor: a module is just a vector space, except that the scalars belong to a ring rather than to a field. But writers of narrative have much more freedom to use metaphors and similes. When you read the words, "After my unfortunate remark, a chill descended on the conversation," you do not need a qualification such as "except that in this case the 'coldness' was in the tones of voices of the people talking and the looks they gave each other."

One could perhaps argue that a great deal of mathematical terminology is already metaphorical. For instance, when we talk of a 26-dimensional space, we are really referring to a mathematical abstraction and not to a real *space*, in the sense of some empty place where one can move around. And many words have associations outside mathematics that have some relation to their mathematical meanings: irrational, function, unbounded, discrete, continuous, converges, differentiate, chaos, contraction, atlas, fiber bundle, foliation, etc., etc. The list is endless.

Another centrally important figure of speech is irony, which I shall define here as writing that is not intended to be taken at face value. If you do not like this definition, it does not matter: it is clearly the case that a great deal of literary writing is not meant to be taken at face value. This looks like a big difference between narrative and mathematical writing. After all, the main objective of the mathematical writer is to explain difficult ideas, and if not everything is supposed to be taken at face value, then that will surely make the task for the poor reader even harder than it already is.

After I spoke at the 2005 “Mathematics and Narrative” conference in Mykonos, a suggestion was made that proofs by contradiction are the mathematician’s version of irony. I’m not sure I agree with that: when we give a proof by contradiction, we make it very clear that we are discussing a counterfactual, so our words *are* intended to be taken at face value. But perhaps this is not necessary. Consider the following passage.

> There are those who believe that every polynomial equation with integer coefficients has a rational solution, a view that leads to some intriguing new ideas. For example, take the equation $x^2 - 2 = 0$. Let p/q be a rational solution. Then $(p/q)^2 - 2 = 0$, from which it follows that $p^2 = 2q^2$. The highest power of 2 that divides p^2 is obviously an even power, since if 2^k is the highest power of 2 that divides p, then 2^{2k} is the highest power of 2 that divides p^2. Similarly, the highest power of 2 that divides $2q^2$ is an odd power, since it is greater by 1 than the highest power that divides q^2. Since p^2 and $2q^2$ are equal, there must exist a positive integer that is both even and odd. Integers with this remarkable property are quite unlike the integers we are familiar with: as such, they are surely worthy of further study.

I find that it conveys the irrationality of $\sqrt{2}$ rather forcefully. But could mathematicians afford to use this literary device? How would a reader be able to tell the difference in intent between what I have just written and the following superficially similar passage?

> There are those who believe that every polynomial equation has a solution, a view that leads to some intriguing new ideas. For example, take the equation $x^2 + 1 = 0$. Let i be a solution of this equation. Then $i^2 + 1 = 0$, from which it follows that $i^2 = -1$. We know that i cannot be positive, since then i^2 would be positive. Similarly, i cannot be negative, since i^2 would again be positive (because the product of two negative numbers is always positive). And i cannot be 0, since $0^2 = 0$. It follows that we have found a number that is not positive, not negative, and not zero. Numbers with this remarkable property are quite unlike the numbers we are familiar with: as such, they are surely worthy of further study.

I admit that I cheated slightly in the second passage by not stressing that the number *i* is introduced as a new kind of number. Nevertheless, even a standard introduction of the complex numbers is highly counterintuitive to many people. If such people want to make progress in understanding mathematics, they normally have to develop a trust that more experienced mathematicians know what they are talking about and mean what they say, however strange what they say might sound. And that makes the use of irony highly problematic. (Occasionally lecturers exploit this as a trick: they present a plausible but incorrect argument as though it is correct, pointing out the fallacy only after they have first persuaded their audience to accept it. This can be a good way of ensuring that the people they are teaching avoid the mistake from then on.)

There is one way in which it would be possible, and potentially a very good idea, for mathematicians to write words that are not true and not intended to be believed. It is not irony, since I am talking not about individual sentences or paragraphs but about an entire genre that we do not have (or if we do, then only to a tiny extent). That is a genre that one might call *mathematics fiction* (by which I do not mean fiction that includes characters who are mathematicians, which we certainly do have), and in particular the subgenre that I think of as fictitious mathematical history. When we learn mathematics, we are typically presented with a highly polished product that is also, as I have already remarked, highly abstract. One way of coming to understand the importance of an abstract definition is to study its history: then one sees what the problems were that the abstract definition helped to solve, what the concrete results were that it helped to synthesize, what the rough edges were that were polished away. However, the *actual* history of an abstract concept is often not by any means the only way of justifying it, or even the best. If it is not the best, then the usual practice is to present a different justification in a purely mathematical way. But such a justification could be far more . . . what is the right word here? . . . ah, yes . . . vivid if it was presented *as though* it had been the reason the concept was first formulated.

The nearest I know to such a piece of fictitious history is a beautiful account by Scott Aaronson of how quantum mechanics could have been discovered. Here are a couple of quotations from a lecture he gave (http://www.scottaaronson.com/democritus/lec9.html).

> There are two ways to teach quantum mechanics. The first way—which for most physicists today is still the only way—follows the historical order in which the ideas were discovered. So, you start with classical mechanics and electrodynamics, solving lots of grueling differential equations at every step. Then you learn about the "blackbody paradox" and various strange experimental results, and the great crisis these things posed for physics. Next you learn a complicated patchwork of ideas that physicists invented between 1900 and 1926 to try to make the crisis go away. Then, if you're lucky, after years of study you

> finally get around to the central conceptual point: that nature is described not by *probabilities* (which are always nonnegative), but by numbers called *amplitudes* that can be positive, negative, or even complex. . . .
>
> The second way to teach quantum mechanics leaves a blow-by-blow account of its discovery to the historians, and instead *starts directly from the conceptual core*—namely, a certain generalization of probability theory to allow minus signs. Once you know what the theory is actually *about*, you can *then* sprinkle in physics to taste, and calculate the spectrum of whatever atom you want. This second approach is the one I'll be following here. . . .
>
> My contention in this lecture is the following: *Quantum mechanics is what you would inevitably come up with if you started from probability theory, and then said, let's try to generalize it so that the numbers we used to call "probabilities" can be negative numbers. As such, the theory could have been invented by mathematicians in the 19th century without any input from experiment. It wasn't, but it could have been.*

What I am suggesting is that somebody should go one step further than Aaronson does here. Why not write a mathematical "short story," or even "novel," all about how quantum mechanics was invented in the nineteenth century without any input from experiment? There are many other mathematical concepts that would lend themselves well to such stories, but this would be a particularly good one.

◆ ◆ ◆

I would like to finish by returning to the main theme of this essay, the use of concrete details to convey abstract thoughts, and illustrate it with a few of my favourite passages from literature. Before I do so, let me point out, in case you have missed it, that up to now I have very consciously followed a show-then-tell policy in this essay. But from now on I would like merely to show.

The first passage is the famous opening to Dickens's *Bleak House*.

> LONDON. Michaelmas Term lately over, and the Lord Chancellor sitting in Lincoln's Inn Hall. Implacable November weather. As much mud in the streets, as if the waters had but newly retired from the face of the earth, and it would not be wonderful to meet a Megalosaurus, forty feet long or so, waddling like an elephantine lizard up Holborn Hill. Smoke lowering down from chimney-pots, making a soft black drizzle, with flakes of soot in it as big as full-grown snow-

flakes—gone into mourning, one might imagine, for the death of the sun. Dogs, undistinguishable in mire. Horses, scarcely better; splashed to their very blinkers. Foot passengers, jostling one another's umbrellas, in a general infection of ill-temper, and losing their foot-hold at street-corners, where tens of thousands of other foot passengers have been slipping and sliding since the day broke (if the day ever broke), adding new deposits to the crust upon crust of mud, sticking at those points tenaciously to the pavement, and accumulating at compound interest.

Fog everywhere. Fog up the river, where it flows among green aits and meadows; fog down the river, where it rolls defiled among the tiers of shipping, and the waterside pollutions of a great (and dirty) city. Fog on the Essex marshes, fog on the Kentish heights. Fog creeping into the cabooses of collier-brigs; fog lying out on the yards, and hovering in the rigging of great ships; fog dropping on the gunwales of barges and small boats. Fog in the eyes and throats of ancient Greenwich pensioners, wheezing by the firesides of their wards; fog in the stem and bowl of the afternoon pipe of the wrathful skipper, down in his close cabin; fog cruelly pinching the toes and fingers of his shivering little 'prentice boy on deck. Chance people on the bridges peeping over the parapets into a nether sky of fog, with fog all round them, as if they were up in a balloon, and hanging in the misty clouds.

Gas looming through the fog in divers places in the streets. much as the sun may, from the spongey fields, be seen to loom by husbandman and ploughboy. Most of the shops lighted two hours before their time—as the gas seems to know, for it has a haggard and unwilling look.

The raw afternoon is rawest, and the dense fog is densest, and the muddy streets are muddiest, near that leaden-headed old obstruction, appropriate ornament for the threshold of a leaden-headed old corporation: Temple Bar. And hard by Temple Bar, in Lincoln's Inn Hall, at the very heart of the fog, sits the Lord High Chancellor in his High Court of Chancery.

Never can there come fog too thick, never can there come mud and mire too deep, to assort with the groping and floundering condition which this High Court of Chancery, most pestilent of hoary sinners, holds, this day, in the sight of heaven and earth.

The next two excerpts are not quite an opening but they come very close to the beginning of Jonathan Franzen's novel *The Corrections*:

The *madness* of an autumn prairie cold front coming through. You could feel it: something terrible was going to happen. The sun low in the sky a minor light, a cooling star. Gust after gust of disorder. Trees restless, temperatures falling, the whole northern religion of things coming to an end. No children in the yards here. Shadows lengthened on yellowing zoysia. Red oaks and pin oaks and swamp white oaks rained acorns on houses with no mortgage. Storm windows shuddered in the empty bedrooms. And the drone and hiccup of a clothes dryer, the nasal contention of a leaf blower, the ripening of local apples in a paper bag, the smell of the gasoline with which Alfred Lambert had cleaned the paintbrush from his morning painting of the wicker love seat. . . .

The anxiety of coupons, in a drawer containing candles in designer autumn colors. The coupons were bundled in a rubber band, and Enid was realizing that their expiration dates (often jauntily circled in red by the manufacturer) lay months and even years in the past: that these hundred-odd coupons, whose total face value exceeded sixty dollars (potentially one hundred twenty dollars at the Chiltsville supermarket that doubled coupons), had all gone bad. Tilex, sixty cents off. Excedrin PM, a dollar off. The dates were not even *close*. The dates were *historical*. The alarm bell had been ringing for *years*.

The next passage is a wonderful mixture of showing and telling, from *The Folding Star*, by Alan Hollinghurst.

I swept the rubbish from an armchair and sat down and still got a piece of Lego up the bum. Why did they do it? Why did this dully charming man, who was already working absurdly to support two children, who got up at six each day to commute to town and was sometimes not home till nine, then go inanely on and sire a third? It must be instinct, nothing rational could explain it—instinct or inattention or else what Edie called polyfilla-progenitiveness: having more children to stop up the gaps in a marriage. I was at the age when I couldn't ignore it; my straight friends married and bred, sometimes remarried and bred again, or just bred regardless. I saw them losing the gift of speech, so used to being interrupted by the demands of the young that they began to interrupt themselves, or to prefer the kind of fretful drivel they had become accustomed to. I saw the huge, humiliating vehicles these studs of the GTi were forced to buy: like streamlined dormobiles, with tiers of baby seats and stacks of the grey plastic crap which seemed inseparable from modern infancy. I saw their doped surrender to domestic muddle, not enough letters on the fridge door to spell anything properly, the chairs covered in yoghurt.

Here is a passage not from fiction but from an essay by the critic Anthony Lane, published in the *New Yorker*, about *The Sound of Music*.

> The Prince Charles Cinema sits in central London, a hundred yards east of Piccadilly, between the Notre Dame dance hall and a row of Chinese restaurants. When it opened, in 1991, the idea was that you could catch new and recent pictures for less than two dollars—a fraction of what they cost around the corner, in the plush movie theatres of Leicester Square. Even now, the Prince Charles has nobly resisted the urge to smarten up; the furnishings are a touching tribute to wartime brown, and the stalls, flouting a rule of theatrical design which has obtained since the fifth century B.C., appear to slope downward toward the back, so that customers in the rear seats can enjoy an uncluttered view of their own knees. The cinema shows three or four films a day; come the weekend, everything explodes. Since August, every Friday evening and Sunday afternoon the program has been the same: "Singalong-a-Sound-of-Music."

To finish this essay I would like to quote from what is considered by some to be the greatest novel ever written. Virginia Woolf famously described it as "One of the few English novels written for grown-up people." That novel is George Eliot's *Middlemarch*. The passage may seem an odd choice because it appears to be a counterexample to the main thesis I have put forward: I cannot deny that it conveys the ideas that it conveys very vividly, and yet it appears to tell more than show. Nevertheless, it still achieves the seemingly magical effect that by the end of the passage we know much more about Mr. Casaubon than we have been told explicitly: the general message is conveyed by means of the telling detail (to use a fortuitously apt phrase). So it is a counterexample that calls for a refinement of the thesis rather than its overthrow. As a bonus, the passage is a supremely good example of the use of irony, and also of metaphor and simile. It could even be said to harbor a warning for mathematicians. That is all the excuse I need to include it, and even to end with it.

> Mr Casaubon, as might be expected, spent a great deal of his time at the Grange in these weeks, and the hindrance which courtship occasioned to the progress of his great work—the Key to all Mythologies—naturally made him look forward the more eagerly to the happy termination of courtship. But he had deliberately incurred the hindrance, having made up his mind that it was now time for him to adorn his life with the graces of female companionship, to irradiate the gloom which fatigue was apt to hang over the intervals of studious labour with the play of female fancy, and to secure in this, his culminating age, the solace of female tendance for his declining years. Hence he determined to abandon himself to the stream of feeling, and perhaps was surprised to find what an exceedingly shallow rill it was. As in droughty regions baptism by immersion could only be performed symbolically, so Mr Casaubon found that sprinkling was the utmost approach to

a plunge which his stream would afford him; and he concluded that the poets had much exaggerated the force of masculine passion. Nevertheless, he observed with pleasure that Miss Brooke showed an ardent submissive affection which promised to fulfil his most agreeable previsions of marriage. It had once or twice crossed his mind that possibly there was some deficiency in Dorothea to account for the moderation of his abandonment; but he was unable to discern the deficiency, or to figure to himself a woman who would have pleased him better; so that there was clearly no reason to fall back upon but the exaggerations of human tradition.