

Proofs and Pictures

James Robert Brown

ABSTRACT

Everyone appreciates a clever mathematical picture, but the prevailing attitude is one of scepticism: diagrams, illustrations, and pictures prove nothing; they are psychologically important and heuristically useful, but only a traditional verbal/symbolic proof provides genuine evidence for a purported theorem. Like some other recent writers (Barwise and Etchemendy [1991]; Shin [1994]; and Giaquinto [1994]) I take a different view and argue, from historical considerations and some striking examples, for a positive evidential role for pictures in mathematics.

- 1 *Bolzano's 'Purely Analytic Proof'*
 - 2 *What did Bolzano do?*
 - 3 *Different theorems, different concepts?*
 - 4 *Inductive mathematics*
 - 5 *Instructive examples*
 - 6 *Representation*
 - 7 *A Kantian objection*
 - 8 *Three analogies*
 - 9 *Are pictures explanatory?*
 - 10 *So why worry?*
-

Mathematicians, like the rest of us, cherish clever ideas; in particular they delight in an ingenious picture. But this appreciation does not overwhelm a prevailing scepticism. After all, a diagram is just a special case and so can't establish a general theorem; even worse, it can be downright misleading. Though not universal, the prevailing attitude is that pictures are really no more than heuristic devices; they are psychologically suggestive, and pedagogically important, but they *prove* nothing. I want to oppose this view and to make a case for pictures having a legitimate role to play as evidence and justification, well beyond a heuristic role. In short, pictures can prove things.¹

1 Bolzano's 'Purely Analytic Proof'

Bernard Bolzano proved the intermediate value theorem. This was early in the nineteenth century, and commentators since typically say two things: first, that

¹ Some other recent authors have taken a similar positive view of pictures. For example, see Barwise and Etchemendy [1991]; Shin [1994]; Giaquinto [1994]. The views in this paper will be developed at greater length in my *Proofs and Pictures: Topics in the Philosophy of Mathematics* (Routledge, forthcoming).

Bolzano's work was initially unappreciated and only later brought to light or rediscovered by others such as Cauchy and Weierstrass; second, that thanks to Bolzano and the others, we now have a *rigorous proof* of the theorem, whereas before we only had a good hunch based on a geometrical diagram.

Typical advocates of this view are the historians Boyer [1949] and Kline [1972] who, respectively, discuss Bolzano in chapters called 'The Rigorous Formulation' and 'The Instillation of Rigour'. It is easy to guess from these titles where their hearts lie and how appreciative their view of Bolzano's efforts might be. Mathematicians hold a similar outlook. Most calculus and analysis texts contain a proof of the intermediate value theorem, and often they have a few casual comments about its significance. Apostol, for example, remarks: 'Bolzano . . . was one of the first to recognize that many "obvious" statements about continuous functions require proof' (Apostol [1967], p. 143). Courant and Robbins, in praising Bolzano, say, 'Here for the first time it was recognized that many apparently obvious statements concerning continuous functions can and must be proved if they are to be used in full generality' (Courant and Robbins [1941], p. 312).

The common attitude towards Bolzano reflects the generally accepted attitude towards proofs and pictures. On this view only proofs give us mathematical knowledge; moreover, proofs are derivations; they are verbal/symbolic entities. Pictures, on the other hand, are psychologically useful, often suggestive, and sometimes downright charming—but they do not provide evidence. When this attitude is brought to bear on the intermediate value theorem, it is perfectly natural to conclude that, until Bolzano, we couldn't really be sure the theorem is true.

Let's look at one of three related theorems (sometimes called the intermediate zero theorem) due to Bolzano [1817].

Theorem: If f is continuous on the interval $[a, b]$ and f changes sign from negative to positive (or vice versa), then there is a c between a and b such that $f(c) = 0$.

Here is a proof which, while not exactly Bolzano's, is in the modern spirit which he created.²

Proof: Assume (with no loss of generality) that $f(a) < 0 < f(b)$. Let $S = \{x: a \leq x \leq b \text{ \& } f(x) < 0\}$. This set is not empty, since a is in it; and it is bounded above by b , so it has a least upper bound, c . There are three possibilities.

(1) $f(c) < 0$. If this is true there is an open interval around c , i.e. $(c - \delta,$

² Bolzano uses concepts like *least upper bound* and *greatest lower bound* which he employs in the following theorem: if a property M does not hold of all values of a variable quantity x but holds of all those which are less than a certain quantity u , then there is always a quantity U , which is the largest of those quantities y which are such that every $x < y$ has property M .

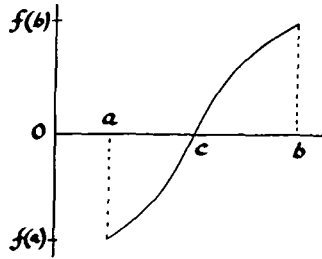


Fig 1. The intermediate zero theorem.

$c + \delta$), in which $f(x) < 0$, for all x in the interval including those greater than c . This contradicts the assumption that c is an upper bound.

- (2) $f(c) > 0$. If this is true there is an open interval around c , i.e. $(c - \delta, c + \delta)$, in which $f(x) > 0$, for all x in the interval, even those less than c . But that's impossible since c is the least of all the upper bounds, so that $f(x) < 0$ for all x less than c .
- (3) $f(c) = 0$: The other two possibilities being ruled out, this one remains. And so, the theorem is proved.

Consider now visual evidence for the theorem. Just look at the picture (Figure 1). We have a continuous line running from below to above x -axis. Clearly, it *must* cross that axis in doing so. Thus understood, it is indeed a 'trivial' and 'obvious' truth.

A simple generalization of this theorem leads to what is now known as the Intermediate Value Theorem, also proved by Bolzano.

Theorem: If f is continuous on the interval $[a, b]$ and there is a C between $f(a)$ and $f(b)$, then there is a c between a and b such that $f(c) = C$.

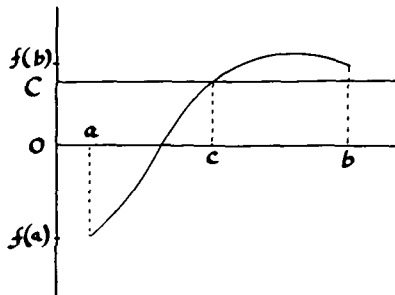


Fig 2. The intermediate value theorem.

I won't bother to give a proof in the Bolzano style, but I will provide another picture (Figure 2).

Bolzano also gives a third theorem, again, a generalization from the others.

Theorem: If f and g are both continuous on the interval $[a, b]$ and $f(a) < g(a)$ and $f(b) > g(b)$, then there is a c between a and b such that $f(c) = g(c)$.

Once we have the hang of the first theorem, we can easily extend the result to the second and third using the same techniques. Again, I shall forgo the analytic proof, but not the visualization. However, this time I'll call on your imagination; like Shakespeare's Prologue on the imagined battlefield of Agincourt, I'll urge you to 'Work your thoughts!' Consider a riddle: a mountain climber starts at the base of a mountain at noon and reaches the top at 6 p.m. She sleeps the night there, then at noon the next day returns to the bottom following the same path. Question: is there a time that afternoon at which she was at the same point on the mountain path both days? The answer, surprisingly, is Yes, in spite of how fast she may go up or down the hill. Here is how to solve the problem: consider an equivalent situation in which we have two hikers, one at the top, the other at the bottom, both setting out at noon on the same day. Obviously, they eventually meet somewhere on the path. And when they do, that is the common time. The solution to the riddle perfectly illustrates the third theorem. It also proves it. Bolzano, of course, gave a 'purely analytic proof', as he called it, not a visualization.

2 What did Bolzano do?

There is a spectrum of ways to understand Bolzano's achievement. The first of these is the common view I mentioned above.

(1) *Bolzano firmly established a theorem that was not known to be true until his proof.* The diagram perhaps played an important heuristic role, but nothing more. Not only is this a common view of the matter, but some of Bolzano's own remarks about the fallibility of geometric intuition strongly suggest that this is how he viewed things.³ But, of course, this is absurd. The geometric picture gives us a very powerful reason for believing the result quite independently of the analytic proof. Using the picture alone, we can be certain of this result—if we can be certain of anything.

Quite aside from the virtues or vices of pictures, we ought to have a somewhat more humble attitude toward our understanding of verbal/symbolic reasoning. First-order logic may be well understood, but what passes for acceptable proof in mathematics includes much more than that. Higher-order logic is commonplace,

³ Years later in his autobiographical sketch he remarked that he had 'found a way to derive from concepts many geometric truths that were known before only on the basis of mere visual appearance' [1969–87], *Vol. 12*, p. 68.

but is far from being house-broken. Moreover, proofs are almost never given in full; they are just sketches which give ample scope for committing some of the well known informal fallacies. Pictures can sometimes even expose verbal fallacies. As for Bolzano in particular, the principles that he used included naïve set theory, later shown to be profoundly inconsistent. A dose of humility seems called for. Consider now a second view of what was achieved.

(II) *Bolzano's proof explained the theorem.* Imre Lakatos [1976] often talks this way about mathematics in general, and would, perhaps, endorse such a view of Bolzano. Philip Kitcher [1975, 1983] holds it explicitly and to some extent so does Alberto Coffa [1991]. The verbal/symbolic proof may well explain the theorem, but the picture explains it, too—at least it explains why the continuous function cuts the x -axis (i.e. at $y = 0$) somewhere or other. The fact that the analytic proof explains the theorem does not set it apart from the picture. (More on explanation below.) Now a third possibility:

(III) *The theorem confirmed the premisses of the proof.* Bolzano is generally considered the 'father of arithmetization', as Felix Klein called him. The arithmetization programme of the nineteenth century sought to found all of analysis on the concepts of arithmetic and to eliminate geometrical notions entirely. (The logicism of Frege and Russell carried this a step further trying to reduce arithmetical notions to logic.) Proving something *already known to be true* was then a feather in the cap of this programme. This method has many champions. Gödel, for example, thought new axioms for set theory should be accepted on the basis of their 'fruitfulness', that is, their good consequences, not their self-evidence. Russell, too, expressed the view clearly: 'we tend to believe the premises because we can see that their consequences are true, instead of believing the consequences because we know the premises to be true' (Russell [1907], pp. 273f.)

It is pretty clear that, of our three options, the final one is the best. (The second option, explanation, is compatible with the third, confirmation, but seems much less plausible.) The consequence of adopting (III) is highly significant for our view of pictures. We can draw the moral quickly: on this view *pictures are crucial*. They provide the *known to be true* consequences that we use for testing the hypothesis of arithmetization. Trying to *get along* without them would be like trying to do theoretical physics without the benefit of experiments to test conjectures.

3 Different theorems, different concepts?

A pair of objections to all this is possible.⁴ One objection is that we have two different proofs of the same result, each with its own strengths and weaknesses.

⁴ Thanks to David Papineau for forcing me to clarify this section (and others).

This I think is quite true—but it is not really an objection. I could rephrase III as saying: the two proof-techniques arrive at the same result. One of these (the picture) is *prima facie* reliable. The other (the analytic proof) is questionable, but our confidence in it is greatly enhanced by the fact that it agrees with the reliable method. This is slightly reminiscent of Whewell's 'consilience of inductions'.

I should add that the way the picture works is much like a direct perception; it is not some sort of encoded argument. However, the boundary between these two ways of understanding the pictures may not be very sharp. Even in fairly simple direct perceptions some 'interpreting' goes on. Ultimately, it may not matter which way we construe the picture, so long as the encoded argument (if there is one) is not the same argument as that given by the verbal/symbolic proof. For, either way, the picture serves as independent evidence for Bolzano's arithmetization programme.

The second possible objection is that we actually have different concepts of continuity at work: one is the ϵ - δ concept, which is more or less Bolzano's; the other is so-called pencil continuity, a geometric notion. To some extent this point seems right; we do have different conceptions of continuity. However, it would be a mistake to infer that the results of the two proofs are incommensurable. For one thing, if they are totally unrelated concepts then it would make no sense to even illustrate Bolzano's theorems with the diagrams, nor would it make any sense to apply Bolzano's result to situations in geometry or mechanics, as is commonly done. If we did take the attitude that these are two quite distinct conceptions of continuity, then we would be very hard pressed to account for a significant amount of mathematical practice. Even if the picture merely does psychological work, that in itself could only be explainable by assuming that ϵ - δ continuity and pencil continuity are somehow deeply related. If they are completely unrelated, then what is the picture doing there? It would be like a dictionary giving a verbal description of apples but illustrating the definition with a picture of a banana.

Perhaps a better understanding of what has historically transpired would be similar to our understanding of what happens in physics or biology. Theories in the natural sciences are tested by observations; however, those very observations are theory-laden. In the act of theorizing about some phenomena we transform the description of the phenomena itself (ducks to rabbits). However, the phenomena—now under a different description—is still relevant for testing. In the case of Bolzano, perhaps the same thing has happened. The concept of continuity has changed, but the diagram is still relevant for testing purposes.

4 Inductive mathematics

It is an uncontentious fact that mathematical reasoning is broader than merely proving theorems. We sometimes forget this when emphasizing the great

achievements of mathematics and the ingenious proofs that have established our most treasured results. But obviously there is more. After all, why work on this problem rather than that one? Why fund this line of research rather than some other? Mathematicians and the mathematical community make all sorts of decisions which are not based on solid analytic proofs. A certain line of research on the Riemann hypothesis is financially supported—not because it is known to be correct, but because it seems promising. Another is rejected as a dead end. Where do these judgements come from? What grounds them? What is the basis of these attitudes which are so crucial to mathematical activity?

Let us call any evidence which falls short of an actual traditional proof, ‘inductive evidence’. Mathematical *achievements* may rest entirely on deductive evidence, but mathematical *practice* is based squarely on the inductive kind. Let’s look briefly at some types.

Enumerative induction: Goldbach’s conjecture says that every even number (greater than 2) is the sum of two primes. Check some examples: $4 = 2 + 2$, $6 = 3 + 3$, $8 = 5 + 3$, $10 = 5 + 5$, $12 = 5 + 7$, and so on. Computers have been used to check this well into the billions. No counter-examples have been found so far. Mathematicians tend to believe that Goldbach’s conjecture is true. They don’t have a proof; but they do have strong inductive evidence.

Analogy: Euler found a way to sum an infinite series that is not ‘rigorous’ by any stretch of the imagination. He argued from analogy that

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}.$$

Polya ([1954], pp. 17ff.) celebrated Euler’s accomplishment, and Putnam [1975] endorsed it, too. Euler’s reasoning was ingenious and persuasive, but not a proof.

Broad experience: Pose a problem; attack it from every conceivable angle; if all plausible approaches lead nowhere, it’s time to think the initial conjecture false. The question, is it true that $P = NP?$, was first posed about twenty five years ago. This is the central problem for those working on computational complexity. There is now a broad consensus in the field that $P \neq NP$. Of course, there is no proof, but a grant proposal which hoped to produce a positive result, would be turned down flat.

These kinds of inductive considerations are central to mathematical activity. Of course, someone could cheerfully grant this sort of thing and then appeal to the traditional distinction between ‘discovery’ and ‘justification’. Inductive evidence, one might claim, plays a role in thinking up theorems, but proofs (and only proofs) give us real justification. (The only difference between this and the distinction philosophers champion is that the distinction here allows the existence of a ‘logic of discovery’ which philosophers often deny.)

But when we turn back to the subject of proofs, we quickly encounter a problem. Analytic proofs, after all, are not constructed *ex nihilo*; they are based on axioms or first principles. But where do these first principles come from? Why do we believe these axioms? Once 'self-evidence' was an acceptable answer, but no more. I mentioned Gödel and Russell above. The Russell passage deserves quoting in full:

we tend to believe the premises because we can see that their consequences are true, instead of believing the consequences because we know the premises to be true. But the inferring of premises from consequences is the essence of induction; thus the method of investigating the principles of mathematics is really an inductive method, and is substantially the same as the method of discovering general laws in any other science (Russell [1907], pp. 273f.)

Gödel shares Russell's consequentialist outlook; that is to say, he too holds that first principles are believed because they have the right consequences, not because they are themselves evident.⁵ But, of course, this view only works because at least *some* of the consequences are evident. Mathematical intuition, as it is often called, must play a role. There are some mathematical truths which are obvious. Gödel and Russell argue that arriving at first principles or axioms in mathematics is similar to science. Mathematical intuitions are like empirical observations in physics. A system of axioms, say for set theory, is postulated just as a theory, say quantum mechanics, is postulated in physics. The theory (in either case) is tested by deriving consequences from it, and is supported by consequences which are intuitive or observational truths, while intuitive or observational falsehoods refute the theory.

The intuitive truths of mathematics need not be certainties any more than ordinary empirical observations must be incorrigible to be confidently used by scientists. The parallel postulate need not be embraced in spite of its intuitive character. And Russell's paradox shows us that some things which seem evident (i.e. that sets exactly correspond to properties) are, in fact, downright false. Still, we can use these intuitions, just as we can use our ordinary eyesight when doing physics, even though we sometimes suffer massive illusions.

The relation for Gödel between a general theory (such as the axioms of set theory) and individual intuitive truths is one of reflective equilibrium, to use a notion introduced by Goodman and made famous by Rawls. That is, we try to construct a theory which is maximally powerful, simple, etc. and which does maximal justice to the intuitive truths. But we allow the possibility that a great mathematical theory will overrule a mathematical intuition, just as a great scientific theory will sometimes overrule an experimental result. The axiom of choice is widely accepted today in spite of some bizarre consequences such as the Tarski–Banach paradox.

⁵ For more on Russell see Irvine [1989].

Even though such famous logicians as Russell and Gödel advocate this view, it has been relatively uninvestigated. Just what is the relation between axioms and intuitions? Should we characterize it as simple H-D? Or perhaps Bayesian? Is Popper’s conjectures and refutations model the right one? Should the intuitive truths be ‘novel’, or can they be already known? These questions have gone largely unexplored. (Lakatos [1976] is a notable exception.)

However, it is not this, but something else in Gödel’s account that I want to focus on, namely the ‘perception’ of mathematical truths. Observational evidence in physics tends to consist in singular spacetime observations: ‘This object, here-now, has property such and such.’ Mathematical intuitions are similar; they are relatively concrete and tend to be singular rather than general (e.g. ‘ $5 + 7 = 12$ ’), though this is certainly not invariable. One thing pictures in particular might do is greatly enlarge the pool of intuitive truths and perhaps even vary their character by adding ones that are relatively more general.

5 Instructive examples

I shall give some example theorems from number theory and infinite series—places where one would least expect to find instructive pictures. In each case the proof will be a diagram. The things to look for are these: is the diagram convincing? Is it a special case (i.e. for some particular n)? But does it establish complete generality? Would a standard verbal/symbolic proof of the theorem, say, by mathematical induction, be more convincing?

Theorem: $1 + 3 + 5 + \dots + (2n - 1) = n^2$

Proof:

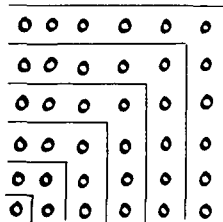


Fig 3.

This picture proof should be contrasted with a traditional proof by mathematical induction which would run as follows:

Proof (traditional): We must show that the formula of the theorem holds for 1 (the basis step), and also that, if it holds for n then it also holds for $n + 1$ (the inductive step).

Basis: $((2 \times 1) - 1) = 1^2$

Inductive: Suppose $1 + 3 + 5 + \dots + (2n - 1) = n^2$ holds as far as n . Now we add the next term in the series, $2(n + 1) - 1$, to each side:

$$1 + 3 + 5 + \dots + (2n - 1) + 2(n + 1) - 1 = n^2 + 2(n + 1) - 1$$

Simplifying the right hand side, we get:

$$\begin{aligned} n^2 + 2(n + 1) - 1 &= n^2 + 2n + 2 - 1 \\ &= n^2 + 2n + 1 \\ &= (n + 1)^2 \end{aligned}$$

This last term has exactly the form we want. And so the theorem is proven.

Theorem:

$$1 + 2 + 3 + \dots + n = \frac{n^2}{2} + \frac{n}{2}$$

Proof:

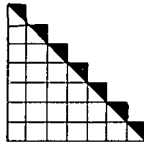


Fig 4.

Again, for the sake of a contrast, here is a traditional proof.

Proof (traditional):

Basis:

$$1 = \frac{1^2}{2} + \frac{1}{2}$$

Inductive:

$$\begin{aligned} 1 + 2 + 3 + \dots + n &= \frac{n^2}{2} + \frac{n}{2} \\ \therefore 1 + 2 + 3 + \dots + n + (n + 1) &= \frac{n^2}{2} + \frac{n}{2} + (n + 1) \\ &= \frac{n^2}{2} + \frac{n}{2} + \frac{2n}{2} + \frac{2}{2} \\ &= \frac{n^2 + 2n + 1}{2} + \frac{n + 1}{2} \\ &= \frac{(n + 1)^2}{2} + \frac{(n + 1)}{2} \end{aligned}$$

Theorem:

$$\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \dots = 1$$

Proof:

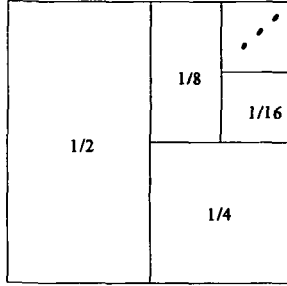


Fig 5.

For the sake of a contrast, here is a standard proof using ϵ - δ techniques:

Proof (traditional): First we note that an infinite series converges to the sum S whenever the sequence of partial sums $\{s_n\}$ converges to S . In this case, the sequence of partial sums is:

$$\begin{aligned} s_1 &= \frac{1}{2} \\ s_2 &= \frac{1}{2} + \frac{1}{4} \\ s_3 &= \frac{1}{2} + \frac{1}{4} + \frac{1}{8} \\ s_4 &= \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} \\ &\dots \end{aligned}$$

The values of these partial sums are:

$$\frac{1}{2}, \frac{3}{4}, \frac{7}{8}, \dots, \frac{2^n - 1}{2^n}.$$

This infinite sequence has the limit 1, provided that for any number ϵ , no matter how small, there is a number $N(\epsilon)$, such that whenever $n > N$, the difference between the general term of the sequence $2^{n-1}/2^n$ and 1 is less than ϵ .

Symbolically,

$$\lim_{n \rightarrow \infty} \frac{2^n - 1}{2^n} = 1 \text{ iff } (\forall \epsilon)(\exists N)n > N \rightarrow \left| 2^n - \frac{1}{2^n} - 1 \right| < \epsilon$$

A bit of algebra gives us the following:

$$\begin{aligned} \left| \frac{2^n - 1}{2^n} - 1 \right| < \epsilon &\rightarrow \left| \frac{-1}{2^n} \right| < \epsilon \\ &\rightarrow 2^n \geq \frac{1}{\epsilon} \\ &\rightarrow \log_2 \frac{1}{\epsilon} \leq n \end{aligned}$$

Thus, we may let

$$N(\epsilon) = \log_2 \frac{1}{\epsilon}$$

Hence,

$$n > \log_2 \frac{1}{\epsilon} \rightarrow \left| \frac{2^n - 1}{2^n} - 1 \right| < \epsilon$$

And so, we have proved that the sum of the series is 1.

Theorem:

$$\sum_{n=1}^{\infty} \frac{1}{4^n} = \frac{1}{4} + \frac{1}{16} + \frac{1}{64} + \frac{1}{256} + \dots = \frac{1}{3}$$

Proof:

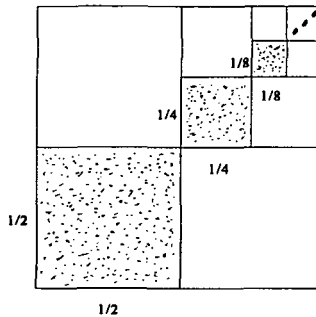


Fig 6.

Again, for comparison's sake, I briefly note the traditional proof:

Proof (traditional): The structure is similar to the one above. We note the sequence of partial sums, the n^{th} being: $s_n = 1/4 + 1/4^2 + 1/4^3 + \dots + 1/4^n$. We next show that

$$\lim_{n \rightarrow \infty} s_n = \frac{1}{3} \text{ iff } (\forall \epsilon)(\exists N)n > N \rightarrow \left| s_n - \frac{1}{3} \right| < \epsilon$$

That completes the proof.

6 Representation

It is probably true that anything can stand for anything. But it is not true that anything can stand *pictorially* for anything. Something special is needed. But what is it about a picture of X that makes it a *picture* of X ? The problem is similar to the problem of intentionality in language and mind: How do words or thoughts get to be *about* things? How do they *represent*? Similarly, how do pictures represent the things they are pictures of?

There is a wide spectrum of views involved in a full answer to these questions, issues involving intentionality, conventions, and so on. I'll avoid most of this and simply focus on a view of pictures which seems highly plausible.

In the *Tractatus*, Wittgenstein made a few cryptic remarks about the relation between pictures and what they picture. 'For a picture to work there must be something in common with what it pictures—"pictorial form"' [Tractatus, 2.161]; 'The minimal commonality between pictorial form and object is *logical form*' [Tractatus, 2.18]. What this suggests is a kind of structural similarity, a notion which is captured by the concept of an *isomorphism*. Barwise and Etchemendy (who are among the very few sympathetic to the use of pictures in inference) explicitly adopt such a view. They hold that 'a good diagram is isomorphic, or at least homomorphic, to the situation it represents' ([1991], p. 22).

In a wide variety of cases, this seems exactly right. It holds, for instance, in the case of the two infinite series examples. But in general it is not so. Consider again the picture proof of the number theory result

$$1 + 2 + 3 \cdots + n = \frac{n^2}{2} + \frac{n}{2}$$

which was given above.

Notice, however, that it is just a picture (in the normal sense) of the $n = 7$ case; and so, we can claim an isomorphism to some number structure with that cardinality. It is certainly not, however, isomorphic to all the numbers. True, it is homomorphic to the whole number structure. But note that a homomorphism to a larger structure is (at least in the case at hand) an isomorphism to a part. The picture (on the Barwise–Etchemendy account) tells us about the isomorphic



Fig 7a,b.

part, but sheds no light on the rest. For example, our picture (Figure 7a) is homomorphic to Figure 7b. But we can't make inferences from the picture (7a) to the non-isomorphic part (in 7b) at all.

There is no useful homomorphism from our picture to all the natural numbers and no isomorphism at all. But still the diagram works. It does much more than establish the formula for $n = 7$; it establishes the result for all numbers. Consequently, I claim, *some 'pictures' are not really pictures, but rather are windows to Plato's heaven*. It is certainly a representation for the $n = 7$ case, but it is not for all generality. For the latter, it works in a different way, more like an instrument. This, of course, is a realist view of mathematics, but not a realist view of pictures. As telescopes help the unaided eye, so some diagrams are instruments (rather than representations) which help the unaided mind's eye.

7 A Kantian objection

Let me quickly try to deal with another potential objection. The diagram (Figure 4 or Figure 7a) that provided the proof of the theorem could be interpreted in a Kantian way.⁶ The claim is this: one sees in the picture the possibility of a reiteration; the diagram can be extended to any number; that's why it works. The objection is anti-platonistic in that it makes a Kantian point about constructability.

This is an interesting objection and I shall say three things in reply. First, my view about pictures is twofold: that they can play an essential role in proofs *and* that there is a platonistic explanation for this. The Kantian objection is only in opposition to the second, platonistic, aspect. Pictures, and seeing the possibility of constructions, can still be a legitimate form of mathematical proof. Indeed, the legitimacy of pictures is upheld by the objection.

Second, the different interpretations of how the picture works are related to the distinction between potential and actual infinities. Both see the formula as holding $\forall n \in \omega$. But the Kantian iteration account sees 4 as a potential infinity only; the platonistic account sees ω as an actual or completed infinity. Of course, the proper understanding of infinity is an unsettled question, but classical mathematics (especially set theory) seems pretty committed to actual infinities. So I see the platonistic interpretation of how the picture works as being favoured for that reason.

Third, we might well wonder: what has the perceived possibility of constructing a diagram of any size got to do with numbers? I certainly don't deny that we can see the possibility of indefinite reiterations of the diagram, but the Kantian objection seems to assume that we know the number theory result

⁶ This was urged by several, especially by W. Demopoulos and R. DiSalle.

because we see the possibility of iteration. I don't know of any argument for this. We could just as well claim that we see the possibility of iteration *because* we see with the mind's eye the number theory result.

8 Three analogies

My main claim—that some pictures are not really representations, but are rather windows to Plato's heaven—will seem highly implausible to most readers. So, beside the discussion so far, I'd like to add three analogies to make things seem a bit more plausible and palatable.

(1) *From aesthetics.* Some with an interest in art and psychology distinguish between a 'pictorial' and a 'symbolic' aspect of a representation (e.g. Arnheim [1969]). Consider a painting such as David's 'Napoleon'. (It depicts Napoleon; he's in a billowing cape, on a spirited white horse, he's pointing ahead.) As a 'picture', it represents Napoleon; as a 'symbol' it represents leadership, courage, adventure. The painting simultaneously manages to be about something concrete and something abstract. It is a wonder that artists can do this, but there is no question that they often do—sometimes brilliantly. (By contrast, my snapshots and doodles are lucky to be pictures.)

What I would like to suggest is that something like this happens with our diagram. It is a *picture* of the special case, $n = 7$, but a *symbol* for every n . Just as we can see courage and adventure depicted in David's painting, so we can see every natural number in the diagram. It's a metaphorical 'seeing', to be sure, but it's the same sort of perception in each case.

(2) *From modern differential geometry.* In modern presentations of differential geometry and general relativity, geometric objects such as: events, vectors pointing from one event to another, tangent vectors, the metric tensor, and so on can be characterized independently of any particular coordinate system. Indeed, it is usually easier and more elegant to do so. But these entities can also be given an explicit coordinate representation, and in practice this is often required. The distance between two events, for example, is an invariant, an objective feature of the geometry; yet its expression in terms of the location of these events can vary greatly, depending on the coordinate system used.

If we consider a surface, the *intrinsic* features are those which characterize the surface independently of any particular coordinatization. By contrast, the *extrinsic* features depend on particular coordinate systems, and change with a change of coordinates. The connection between them is this: an intrinsic feature corresponds to the existence of a coordinate system with specific appropriate extrinsic features.⁷

⁷ Excellent discussions can be found in Aleksandrov [1963] and Friedman [1983].

The distinction suggests something objective about the intrinsic aspects of a surface, and something less so in the extrinsic, due to the arbitrariness of the various forms of representation. Some concerned with spacetime (e.g. Friedman [1983]) argue that this distinction corresponds to a distinction between the factual and the conventional; the intrinsic features of spacetime (curvature, metric tensor, etc.) are objectively real while extrinsic features are mere artefacts of the form of representation.

The analogy with picture-proofs that I want to suggest is this: any representation of a surface, say, displaying its curvature, will always be in some particular coordinate system. Analogously, a picture of a numerical relation will always be with some particular number, n . But an intrinsic feature, such as Gaussian curvature, is independent of any particular coordinate system. Similarly, the evidential relation in the number theory diagram is independent of any particular n -element picture. To calculate the Gaussian curvature, however, some particular coordinate system is required. Similarly, to grasp the evidential relation some particular n -element picture is needed.

Because of the analogy I am tempted to call number theory diagrams 'extrinsic pictures', since they are particular representations like particular coordinate systems. Is there such a thing as an intrinsic number theory diagram? Of course. It is the one seen by the mind's eye—and it has no particular number of elements in it. That is the one we grasp, the one that provides the evidence for the theorem.

(3) *From 'Natural Kind' reasoning*: One sort of inductive inference is 'enumerative'. We notice that all of a very large number of observed ravens are black; so we infer that *all* ravens are black. Often in science, powerful inferences are made from a single case. In high-energy physics, for instance, a single event (sometimes called a 'golden event') captured in a bubble chamber photo will sometimes be sufficient evidence for a powerful general conclusion. One positron is sufficient to generalize about the mass of all others. One sample of water is sufficient to establish that all water is H₂O. The form of inference seems based on a principle that runs like this: *if X is a natural kind and has essential property P, then all instances of the kind have property P*. The assumption at work would then be that positrons or water are natural kinds and that their mass or chemical composition are essential properties. In principle, only one instance is needed to allow us to draw the general conclusion about all positrons or all water. In practice, a few more than one instance are likely to be necessary, simply to build confidence that no mistakes were made in measuring the mass or analysing the composition of the sample. Even so, the power of natural kind inference is remarkable.

Something analogous to natural kind inference is going on in the number theory picture-proof. We can take the diagram $n = 7$ to be an instance of a

natural kind; and we further take the formula $n^2/2 + n/2 = 1 + 2 + 3 + \dots + n$, which is true in this case, to be a natural kind property. Since it holds for $n = 7$, it holds for all n .

The fly in the ointment is this: water is *essentially* H_2O , but only *accidentally* thirst-quenching. Mathematical objects would seem to have only essential properties, and it would be a terrible inference to pass from ‘7 is prime’, thus all numbers are prime. So, I am reluctant to see the inference from a picture of a special case to all generality as a clear case of natural kind inference; but it is interestingly similar.

9 Are pictures explanatory?

Mathematicians look for two things in a proof—evidence and insight. Traditionally, a proof must firmly establish the theorem. That, for just about everyone except Lakatos, is a *sine qua non* for any proof. But a good proof also helps us to understand what’s going on. Insight, understanding, explanation, are somewhat nebulous, but highly desirable. Proofs needn’t have them, but are cherished when they do.

Are picture proofs rich in insight? Many commentators suggest as much; they even play evidence and insight off against one another, suggesting that what we lose in rigour we make up in understanding. (Polya is perhaps the best example of this.) However, this seems slightly misguided. And I would be seriously misunderstood, if it was thought I am suggesting that it is worth giving up some rigour in exchange for insight. This is doubly wrong. I don’t see any abandoning of rigour by allowing the legitimacy of picture proofs. And second, greater insight isn’t always to be found in pictures.

In the two number theory cases above, a proof by induction is probably more insightful and explanatory than the picture proofs. I suspect that induction—the passage from n to $n + 1$ —more than any other feature, best characterizes the natural numbers. That’s why a standard proof is in many ways better—it is more explanatory.

To be sure, some insight is garnered from the diagrams which prove the two infinite series examples. From looking at them, we understand why the series have the sums that they do. Pictures often yield insight, but that is not essential. They are mainly a form of evidence—a different form, to be sure, than verbal/symbolic proofs; but they have the same ability to provide justification, sometimes with and sometimes without the bonus of insight and understanding.

10 So why worry?

Philosophers and mathematicians have long worried about diagrams in mathematical reasoning—and rightly so; they can indeed be highly misleading.

Anyone who has studied mathematics in the usual way has seen lots of examples that fly in the face of reasonable expectations. I'll give one now which is not so well known, but highly striking.

Draw four circles in the plane, centred at $(\pm 1, \pm 1)$, each with radius 1. (Figure 8a.) Draw a fifth circle, this time at the origin, so that it touches each of the four circles. Draw a box around the four circles. It will have sides stretching from -2 to $+2$. Obviously, the inner circle is completely contained within the box. Do the same in three-dimensional space, this time drawing eight spheres centred at $(\pm 1, \pm 1, \pm 1)$ and a ninth sphere at the origin touching the other eight. Draw a box around the eight spheres. Once again, the central sphere is completely contained within the box. (Figure 8b.)

Reflecting on these pictures, it would be perfectly reasonable to jump to the 'obvious' conclusion that the result holds in higher dimensions. Amazingly, this is not so. At ten dimensions or higher, the central sphere breaks through the n -dimensional box. Here's why. The distance from the origin to the centre of any sphere is $\sqrt{((\pm 1)^2 + \dots + (\pm 1)^2)} = \sqrt{n}$. But each sphere has radius 1; thus, the radius of the central sphere is $\sqrt{n} - 1$. For $n \geq 10$, we have $\sqrt{n} - 1 > 2$. Thus, the central sphere will break through the sides of the n -dimensional box—a profound shock to intuition.

What is the moral to be drawn from examples such as this? The all too common response is to relegate pictures to heuristic status only; they are not to be trusted as serious mathematical evidence. But as I pointed out above, verbal/symbolic proofs can mislead, too. Pictures are no worse, and can even correct faulty derivations. It would be much better to consider the evidence acquired from pictures to be like the empirical evidence acquired from microscopes, bubble chambers, and other instruments for making observations. These instruments can be highly misleading, too. Optical properties and staining techniques which were not understood have led microscope users to 'observe' things that are not real but that were merely artefacts of the process of observation.

It would be silly to tell people: *just be careful*. We have to learn how microscopes and spark chambers, etc. work. As we learn, the quality of our

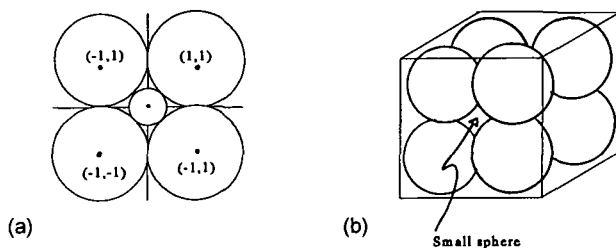


Fig 8a,b.

observations improves. Learning about the instruments and learning about nature go hand in hand. The same can be said about pictures in mathematics. The fact that many mislead is no reason to reject them in principle as a source of evidence. We simply *have to learn* how to use them, just as we must continue to learn more about microscopes. This is a process which will never end.

I realize that talk of ‘the mind’s eye’ and ‘seeing mathematical entities’ is highly metaphorical. This is to be regretted—but not repented. Picture proofs are obviously too effective to be dismissed and they are potentially too powerful to be ignored. Making sympathetic sense of them is what is required of us.

Acknowledgements

I am grateful to the members of several audiences where I have delivered versions of this paper, to David Papineau and the referees of this journal for many useful comments, and also to SSHRC for its support.

*Department of Philosophy
University of Toronto
Toronto M5S 1A1
Canada*

References

- Aleksandrov, A. D. [1963]: ‘Curves and Surfaces’, in A.D. Aleksandrov *et al.* (eds), *Mathematics: Its Content, Method, and Meaning*, Cambridge, MA, MIT Press.
- Apostol, T. [1967]: *Calculus*, Waltham, MA, Blaisdell.
- Arnheim, R. [1969]: *Visual Thinking*, Berkeley, CA, University of California Press.
- Barwise, J. and Etchemendy, J. [1991]: ‘Visual Information and Valid Reasoning’, in W. Zimmerman and S. Cunningham (eds), *Visualization in Teaching and Learning Mathematics*, Mathematical Association of America.
- Bolzano, B. [1817]: *Rein analytischer Beweis des Lehrsatzes, dass zwischen je zwei Werthen, die ein entgegengesetztes Resultat gewähren, wenigstens eine reele Wurzel der Gleichung liege*, Prague, Gottlieb Hass.
- Bolzano, B. [1851]: *Paradoxes of the Infinite* (trans. D. A. Steele), London, Routledge & Kegan Paul [1950].
- Bolzano, B. [1969/87]: *Gesamtausgabe, Vols. 1–15*. Stuttgart, Fromman.
- Boyer, C. B. [1949]: *The Concepts of the Calculus*, New York, Dover (reprint).
- Brown, J. R. [1990]: ‘II in the Sky’, in Irvine (ed.) [1990].
- Brown, J. R. [1991]: *The Laboratory of the Mind: Thought Experiments in the Natural Sciences*, London and New York, Routledge.
- Brown, J. R. [1994]: *Smoke and Mirrors: How Science Reflects Reality*, London and New York, Routledge.

- Coffa, J. A. [1991]: *The Semantic Tradition from Kant to Carnap*, Cambridge, Cambridge University Press.
- Courant, R. and Robbins, H. [1941]: *What is Mathematics?* Oxford, Oxford University Press.
- Friedman, M. [1983]: *Foundations of Spacetime Theories*, Princeton, Princeton University Press.
- Giaquinto, M. [1994]: 'Epistemology of Visual Thinking in Elementary Real Analysis', *British Journal for the Philosophy of Science*, **45**, pp. 789–813.
- Gödel, K. [1944]: 'Russell's Mathematical Logic', reprinted in P. Benacerraf and H. Putnam (eds), *Philosophy of Mathematics*, Cambridge, Cambridge University Press.
- Gödel, K. [1947]: 'What is Cantor's Continuum Problem?', reprinted in P. Benacerraf and H. Putnam (eds), *Philosophy of Mathematics*, Cambridge, Cambridge University Press.
- Goodman, N. [1976]: *Languages of Art*, Indianapolis, Hackett.
- Irvine, A. [1989]: 'Epistemic Logicism and Russell's Regressive Method', *Philosophical Studies*, **55**, pp. 303–27.
- Irvine, A. (ed.) [1990]: *Physicalism in Mathematics*, Dordrecht, Kluwer.
- Kitcher, P. [1975]: 'Bolzano's Ideal of Algebraic Analysis', *Studies in the History and Philosophy of Science*, **6**, pp. 229–71.
- Kitcher, P. [1983]: *The Nature of Mathematical Knowledge*, Oxford, Oxford University Press.
- Lakatos, I. [1976]: *Proofs and Refutations*, Cambridge, Cambridge University Press.
- Kline, M. [1972]: *Mathematical Thought From Ancient to Modern Times*, Oxford, Oxford University Press.
- Polya, G. [1954]: *Mathematics and Plausible Reasoning*, Princeton, Princeton University Press
- Putnam, H. [1975]: 'What is Mathematical Truth?' *Mathematics, Matter and Method: Philosophical Papers, Vol. 1*, Cambridge, Cambridge University Press.
- Russell, B. [1907]: 'The Regressive Method of Discovering the Premises of Mathematics', reprinted in *Essays in Analysis* (ed. by D. Lacky), New York, George Braziller.
- Shin, Sun-Joo [1994]: *The Logical Status of Diagrams*, Cambridge, Cambridge University Press.
- Wittgenstein, L. [1921]: *Tractatus Logico-Philosophicus*, London, Routledge.