# Mathematical arguments in context

Jean Paul Van Bendegem and Bart Van Kerkhove*
*Centre for Logic and Philosophy of Science, Vrije Universiteit Brussel*

**Abstract.** Except in very poor mathematical contexts, mathematical arguments do not stand in isolation of other mathematical arguments. Rather, they form trains of formal and informal arguments, adding up to interconnected theorems, theories and eventually entire fields. This paper critically comments on some common views on the relation between formal and informal mathematical arguments, most particularly applications of Toulmin's argumentation model, and launches a number of alternative ideas of presentation inviting the contextualization of pieces of mathematical reasoning within encompassing bodies of explicit and implicit, formal and informal background knowledge.

**Keywords:** contextual knowledge, mathematical arguments, mathematical practice, Toulmin, rhetoric

## 1. Introduction

The general claim behind this paper, of which it is our (modest) aim to undertake a first, bold exploration, is that mathematical arguments, formal or not, need to be put in their proper context as much as any other piece of human knowledge. In order to avoid misunderstanding, we would like to stress a point implicit in the latter claim, as it is one of the core ideas behind our approach: both formal and informal pieces of mathematical reasoning in our view constitute mathematical arguments, not just the informal ones. We begin, in section 2, by distinguishing a number of common views on the relation between formal mathematical arguments, i.e. proofs, and informal mathematical arguments. One of those, the view that formal proofs constitute a special kind of argument, is central to applications of Toulmin's model of a general argument to the case of mathematics. In section 3, we point to what we think are shortcomings of the latter style of argumentative analysis. The most important of these is that it does not facilitate or invite the contextualization of particular pieces of mathematical reasoning within the vast body of explicit and implicit, formal and informal background knowledge. Sections 4 and 5 present a number of illustrative cases in this respect, along the lines of what Philip J. Davis once called skipping (of steps) and splicing (of tasks) in mathematical

* The latter author acknowledges support by the Alexander von Humboldt Foundation, Germany, to which he has been affiliated.

argumentation. In section 6, what we have thus developed is connected to the work of Reviel Netz, who is among those to suggest that the possibility of alternative modes of mathematical presentation, within a large body of interconnected mathematical knowledge, reveals the irreducible rhetorical nature of mathematics.

## 2. Views on proofs *versus* arguments

Probably the most common view concerning the interaction between mathematical proofs and (informal) arguments, is that mathematics solely relies on formal proofs, and that any other arguments are to be reserved for the non-mathematical sciences and other areas of human knowledge. Mathematics has a special status. In mathematical discourse, there is no room for non-formal arguments, only for undisputed moves: unclarity is suspended and transparency reigns. Thus severing mathematical from everyday reasoning, mathematics is established as a very particular language game. Because of this extreme separation, for the philosopher, nothing worth discussing remains. We therefore quickly move on to more fruitful discussions.

The second most common view is that mathematics is indeed about formal proofs, but that informal arguments can nevertheless play a role in it. Inductive, probabilistic, computerized, visual, intuitive, analogical or metaphorical modes of reasoning are among the candidates. Whether the actual role played by these techniques is or can be an epistemically important or even central one, is another issue. For example, while the pedagogical or illustrative value of pictures or diagrams is virtually unquestioned, the matter of whether they may also be the heuristic driving force behind mathematical discovery, both in its conjecturing and proving phases, remains highly controversial.[1] This said, it is a fact that almost no mathematical proofs satisfy logical standards in their entirety. As a result, exactly what pieces of mathematical reasoning are to be considered 'rigorous' and thus acceptable, is a matter itself subject to (non-formal) argumentation. In line with this, also on an argumentative level are any decisions, to be reached within the mathematical community, about what are the important, interesting and/or relevant problems to be pursued. Like a number of aforementioned proof methods (save the mechanical ones), the latter topics lie within the sphere of understanding and explanation, where what is at stake is not *whether* but rather *why* propositions and conjectures are true.

[1] Readers keen on putting their faith to the test, might want to consult the dozens of cases presented in Nelsen (1993) and (2000).

Let us start making the transition to another, more challenging view. When looked upon at the level of an entire proof, the idea can be launched that — as such a proof rarely seems to explain anything at all — a strictly formal proof needs to be turned into a convincing argument, functioning as its complement. Consider the following theorem. The function $f$ from $I\!N \times I\!N$ to $I\!N$: $(a,b) \rightarrow (a+b)(a+b+1)/2+a = f(a,b)$ is a bijection. A purely mathematical proof proceeds as follows. Prove that if $f(a,b) = f(c,d)$ then $(a,b) = (c,d)$ or $a = c$ and $b = d$. Although this looks tremendously ugly, it can indeed be done. An entirely different proof, which is informal in nature and very hard to express mathematically, but on the other hand has the virtue of being absolutely clear, would read as follows. Define an order relation $<$, such that: $(c,d) < (a,b)$ iff $c+d < a+b$, or $(c+d = a+b$ and $c < a)$. Using this order relation, determine how many couples precede some given $(a,b)$, where $a+b = n$. If $c+d \neq a+b$ then $c+d < n$, so $c+d$ can take $n$ values: $0,1,2,\ldots,n-1$. For every value $i$, $i+1$ couples $(0,i),\ldots,(i,0)$ are possible. This makes $1+2+3+\ldots+n = n(n+1)/2$ or $(a+b)(a+b+1)/2$, to which are added the couples that satisfy $c+d = n$ and $c < a$, to obtain $f(a,b)$ couples preceding $(a,b)$. So $f$ is a bijection.[2]

Drawing upon the former paragraph, then, a third, quite untraditional view on the formal *versus* the informal can be discerned: a mathematical proof, on the whole, is a kind of argument. This indeed appears to be the idea behind recent applications to the case of mathematics of Toulmin's popular six-fold model of a general argument (see, e.g., Aberdein, 2007). Attention is drawn here not to the formal validity of an argument, but rather to its overall structure. What is thus laid bare, is, first, the general outline of a proof, or a sketch of its characteristic global shape, which when properly done adds to its surveyability and robustness, and, second, an identification of the different types of warrants or sub-arguments used in the course of the proof. It is clear that an exercise like this does not yield a formal, but

[2] Mathematical conviction can be a tricky affair though, exhibiting a high degree of counter-intuitiveness *vis-à-vis* everyday reasoning. Take the following, classic problem. Today is day 1, and you receive 1 euro. Tomorrow, day 2, you receive $\frac{1}{2}$ euro, day 3, you get $\frac{1}{3}$ euro, and so on. (We make a slight abstraction here, assuming all sums of money actually exist.) Generally speaking, on the $n$th day, $\frac{1}{n}$ euro is given to you. Finally, assume that this process continues *ad infinitum*, until the end of times. Question: How rich will you be? We guess that most, if not all, *non*-mathematicians will wonder whether the term rich should not be put between quotation marks, as in the real world this clearly is *not* a successful road to wealth. And yet, as mathematicians know, at the end of times, one will unavoidably be filthy rich, having at disposal an infinite amount of money.

instead an informal view into (mathematical) reasoning. Next some considerations about this particular view.

## 3. About Toulmin-style analyses

One of the questions popping up in the course of applying an analytic tool in the vein of Toulmin's is what needs to be shared to let a mathematical proof under examination function as an argument. In our view, this goal is not properly served by subsequently comparing one proof and one (schematized) argument, as is indeed done in Toulmin-style analyses.[3] This is so in general, because the stress laid on formal-informal contrasts obscures rather than clarifies the goal, viz. to show that mathematical proofs are arguments after all. But there are a number of more specific weaknesses as well, which have to do with the potential neglect of factors that are of utmost relevance to (a philosophical assessment of) mathematical arguing. First, philosophical comparisons between deductive and (also mutually differing) non-deductive kinds of warrants and qualifiers (see above: induction, probability, computing, vision, intuition, analogy, metaphor) are hardly facilitated. Take the case of analogous reasoning, which seems difficult to put one's finger on. George Pólya, in his 1954 book *Mathematics and plausible reasoning*, has defined it as a type of conceptual similarity essentially rooted in the intentions of the thinker drawing it, leaving clarification of analogy to consist in pinpointing the shared concepts involved.[4]

As an illustration, Pólya (ch.2, §6) develops the historical example of Euler's derivation, by analogy, of a formula expressing the sum of the reciprocals of the subsequent squares $1 + \frac{1}{4} + \frac{1}{9} + \frac{1}{16} + \ldots$, viz., $\frac{\pi^2}{6}$. Euler arrived at it by applying a decomposition rule of algebraic equations to an equation which is clearly not of that type, viz., $\sin x = 0$ (being of infinite degree). This was a daring move, but Euler found he had good reasons to be confident, as "the numerical value for the sum of the series which he [had] computed before, agreed to the last place with $\frac{\pi^2}{6}$" (Pólya, 1973, p. 20). Notice that these reasons were inductive in their turn, and not demonstrative, although later, Euler was also able to actually prove the equality, and thus formally confirm his informal reasoning. While undoubtedly in favour of the exact mode

[3] As has been pointed out to us, Toulmin-style exercises in representing more complex mathematical arguments or multi-step proofs are available. We get back to this below.

[4] E.g., two systems are analogous "if they agree in clearly definable relations of their respective parts" (Pólya, 1973, p. 13).

of mathematical reasoning (like Pólya), "Euler seems to think the same way as reasonable people, scientists or non-scientists, usually think. He seems to accept certain principles: *A conjecture becomes more credible by the verification of any new consequence*. And: *A conjecture becomes more credible if an analogous conjecture becomes more credible*" (Pólya, 1973, p. 22; original emphasis). This said, induction alone clearly does not suffice. Typically, a ('true'?) mathematician cannot rest content with it, and wants to pass from the inductive into the demonstrative phase of inquiry. "The verification of any consequence increases our confidence in the conjecture, but the verification of the [inductive generalization] can do more: it can *prove* the conjecture" (Pólya, 1973, p. 110). Here the affinity between the several "alternative" modes of mathematical reasoning listed above becomes apparent. Our goal is to explore and define the border between formal or deductive proof and their non-formal and non-deductive complements.[5]

Let it be clear that we do not want to underestimate the power of Toulmin-style (or similar) schemes.[6] Abstracting (for now) from a possible conflation of what Toulmin distinguishes as 'warrant-using' and 'warrant-establishing' arguments, one mathematical, the other metamathematical,[7] it might not even be a very difficult exercise to put the Euler-argument in such a schema (see figure 1). The datum D is the 'proof' or reasoning, the claim C is the statement that the sum of the inverse squares is indeed $\frac{\pi^2}{6}$ and the warrant W should then be something like "If the reasoning can be transformed into a formal or 'decent' proof." In the extended scheme the warrant can be supported by a backing B that refers either to the mathematician's intuition or to the brute fact of calculating both expressions, something that Euler actually did. The qualifier Q is to be read as 'presumably' or 'hopefully' and expresses doubts about this transformation, and the rebuttal R amounts to, e.g., a proof that shows that the reasoning cannot be transformed into a 'decent' proof. As said, such a translation into the Toulmin-scheme proved its force, but it does tend to incorporate the non-formal arguments into a single framework with the formal arguments. In other words, all informal elements are seen as related to the formal elements and not necessarily as elements on their own.

[5] For a consideration about inductive and probability arguments in this respect, e.g., we refer to Van Kerkhove and Van Bendegem (2008). For an exploration of an even more daring idea, the possibility of mathematical rhetoric, see Van Bendegem ([to appear I]).

[6] Alternative approaches could include argumentation schemes or dialogical logic, as was suggested by one of the referees.

[7] We thank Andrew Aberdein for pointing this out to us.

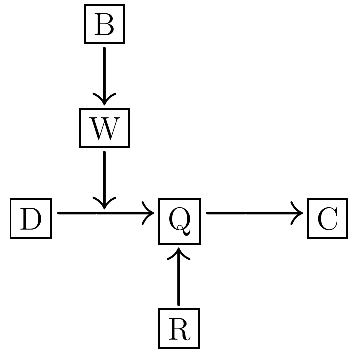


*Figure 1.* Basic Toulmin-style scheme

Another illustration is the complexity of a proof itself. That the Toulmin scheme allows for the possibility of 'embedding,' as is shown in Aberdein (2006), and of 'folding,' where the proof is in a sense summarized by converting the embeddings in single steps, is surely an important technique for dealing with such complexities. However, we believe it can at the same time be considered to be another apparent weakness of this type of analysis of mathematical arguments, because all levels are treated in a similar or homogeneous way. Imagine an embedding with several levels. If one is looking at what happens at a certain level, this need not necessarily imply that one knows the precise level one is looking at.[8]

As will become clear in the sequel, when dealing with extremely long or complicated proofs, additional features are needed, such as the possibility to accommodate *and differentiate* several plots and sub-plots. Consider, in this respect, two implicitly but widely applied techniques potentially giving rise to mathematical alienation, as distinguished by Davis (1972): *splicing* and *skipping*. Both are aimed at cutting down the size of formal proofs, in order to improve their surveyability. That is, steps in the proof are dropped, either because they are extensively argued for elsewhere (splicing), or because they are simply thought to be intuitively clear or acceptable (skipping). In the latter case, sometimes no formal argument has even been produced within the mathematical community, and grasping the move is supposed to be an irreducible feature of this particular community's membership, i.e. of being famil-

[8] As for global pictures of mathematical proofs (or, at a higher level, of theories or subdisciplines), these for us do not belong to the so-called rhetorical domain, where one convinces rather than proves. Instead, together with and as a complement of the formal development, they are part and parcel of mathematical proof. This latter point also implies that, in their turn, the outlines or global structures are in no way more fundamental, but are in as much need of their formal counterparts than the other way round.

iar enough with its central body of knowledge to 'see' why the step is acceptable.[9] A perfect example is the famous pigeonhole principle (PHP). For most mathematicians PHP does not require proof: what needs to be shown about the fact that, if you have $n$ objects and less than $n$ boxes, then one box must contain at least 2 objects? However, in foundational work proofs of PHP do exist, but that does not seem to bother the working mathematician. No better illustration than the contrast between, on the one hand, the words of Aigner and Ziegler (1998, p. 123) about PHP — "Well, this is indeed obvious, there is nothing to prove" — and Kolodziejczyk and Thapen (2008), a recent paper in *The Journal of Symbolic Logic* actually examining particular models where (a weak version of) PHP fails!

We shall further address the matter of plots and sub-plots in sections 5 and 6. First, however, in section 4, we give an example of the former of the two techniques mentioned above, viz. splicing.

## 4. The Enormous Theorem

A special case of splicing is the division of labour, geographically as well as in time, and either intentionally or not, when constructing a proof. A magnificent opportunity for illustrating it, and the expansive nature of mathematical theories in general, is offered by the project of classifying finite simple groups. Initiated by Otto Hölder (1859–1937) in his paper *Die einfachen Gruppen in ersten und zweiten Hundert der Ordnungszahlen* (The simple groups of orders up to one and two hundred; 1892), on the basis of foregoing work by Galois and Jordan among others (see Kline, 1990, §31.6), the history of this project spanned the entire twentieth century. The *Classification Theorem of Finite Simple Groups*, also known as the Enormous Theorem, states that all simple groups of finite order can be classified into one of four types. Three of these cover infinitely many cases, viz., *cyclic* groups $Z_p$ (of prime order), *alternating* groups $A_{n\geq 5}$ (of degree at least five), and *Lie-type* (twisted) Chevalley or Tits groups, the fourth is a 'residual category' of 26 *sporadic* groups. The first serious attempts at completing a proof were made in the 1950s, but it was not until the 1960s that any breakthroughs were reported. From the mid 1970s, it became gradually clear

[9] Compare Thurston, 1994, p. 168: "Within any field, there are certain theorems and certain techniques that are generally known and generally accepted. When you write a paper, you refer to these without proof. [...] Many of the things that are generally known are things for which there may be no known written source. As long as people in the field are comfortable that the idea works, it doesn't need to have a formal written source."

that the supply of sporadic simple groups had come near to exhaustion. Although a last one was apparently still left to be constructed, viz., the famous *Monster*, or the highest order sporadic group, by 1980 theorists felt confident enough to declare the problem solved.[10]

> The proof, however, was unlike anything ever seen before in mathematics. Most proofs can be read and verified by anyone knowledgeable in the field. But most proofs are at most a few pages long, and those that aren't self-contained rest either on a well-developed theory or on a handful of references. None of that held for the Enormous Theorem, and no one had the time, the patience, and the expertise to double-check every nook and cranny of the myriad papers that comprised its proof. Furthermore, researchers acknowledged, many of the individual papers almost certainly contained errors — usually bad news for a mathematical proof. (Cipra, 1996, p. 87)

The reason for all this: the purported proof, in its entirety, consisted of some fifteen thousand journal pages, spread over hundreds of articles (some of them still unpublished), written by dozens of authors. It was the concluding result of the Classification Programme launched and coordinated by Daniel Gorenstein (1923–1992) from the early 1970s. The project proposed a type of strategy for joining forces in pinpointing finite simple groups that was first laid out in the early 1950s, viz., by Richard Brauer. Gorenstein therefore appropriately called the period 1950–1980 the 'Thirty Years War,' a conflict to whose decision he contributed greatly. "Not a single leading group theorist besides Gorenstein believed in 1972 that the Classification would be completed this century. By 1976, almost everyone believed that the Classification problem was busted" (Solomon, 1995, p. 235). However, from this moment, increasing paper output began to give rise to problems of surveyability. "At least 3,000 pages of mathematically dense preprints appeared in the years 1976-1980 and simply overwhelmed the digestive system of the group theory community" (Solomon, 1995, p. 236). For example, the preprint *Quasithin finite groups* by Geoff Mason, part and parcel of the Enormous Theorem, remained incompletely assimilated, unpublished, and thus controversial throughout the 1980s. Moreover, the reluctance turned out to be justified. "Around 1989, [Michael] Aschbacher noticed that Mason's 800-page manuscript on quasithin groups was incomplete in various ways; in particular it lacked a treatment of certain 'small' cases" (Solomon, 2001, p. 341). Aschbacher then set out to fill the gap and propose an alternative proof of the Quasithin Theorem, together

[10] See Solomon (2001) for an excellent historical overview of the major contributions.

with Stephen Smith, a task which has been completed only recently.[11] The general opinion now seems to be that the Enormous Theorem has hereby finally been established.

As we hope the former paragraphs have shown, if one considers mathematical proofs as arguments, to adequately picture them is not at all a straightforward job. In the rest of the paper, inspired by what Philip Davis has called *skipping* (see the end of section 3), we would like to further extend or complicate the standard application of the models of argumentation theory. More particularly, we want to illustrate that a mathematical proof is a type of argument that requires a considerable amount of formal and informal *background knowledge and training* in order to be put to fruitful use. From an application perspective this seems obvious, for when a piece of mathematics is applied, no one discusses the correctness of the proofs, but rather one argues about the appropriateness of the application itself. This means that in such a case the mathematical proof is not *de facto* stronger than an argument showing the application to be incorrect. In what follows, we want to focus on a similar kind of dynamic *within* research mathematics.

## 5. Pólya and some conjectures in number theory

In this section, it is our aim to further illustrate previous considerations in terms of an analysis of (part of) a specific mathematical paper. This will not be done merely in terms of how the mathematical proofs are presented, but also (and especially) in terms of the way theorems and proofs are organized in the paper, i.e. of its global argumentational structure. The text we have chosen is a short paper by George Pólya, a mere nine pages in published form, entitled *Verschiedene Bemerkungen zur Zahlentheorie* (Several remarks on number theory). Apart from being a proper tribute to the mathematical humanist Pólya, there is nothing particular about this choice, as it is our claim that virtually any mathematical text will do.[12]

[11] The result, a proof about 1,200 pages long, was published by the AMS in 2004 under the title *The Classification of Quasithin Groups* (two volumes). For a summary, see Aschbacher, 2004.

[12] For the *petite histoire*, the former author of this paper, many years ago, read a piece in a book of essays by the French mathematician and writer Raymond Queneau, in which false conjectures in number theory were dealt with (Queneau 1978). One of these was proposed by Pólya, which led to the original paper, discovering along the way that Queneau had misunderstood the conjecture and in fact had considered it to be the opposite statement. Reading Pólya's text made clear this was a prime candidate for an argumentative-rhetorical analysis.

The procedure we follow is quite simple: we shall go through (part of) the paper, step by step, paying attention to every detail, asking over and over how each move contributes to the overall structure of the text. It is — and it is our suggestion that one should expect this — a process that is closely related and similar to what goes on in the mathematician's mind when reading a paper. Seen from this angle, we are basically thinking out loud. Note that we are not incorporating actual fragments of the paper itself, for reasons of available space, but also because the paper is very easily accessible.[13]

On to the paper itself then, which consist of three parts. We focus on the opening paragraphs of the first of these, entitled *Über die Irreduzibilität gewisser ganzzahliger Polynome* (On the irreducibility of certain integer polynomials). The paragraph in question starts with a definition and an observation. The definition concerns polynomials with integer coefficients, such as $f(x) = 5x^2 + 8x - 10$, and the values of $f(x)$, $x$ also being an integer. The observation is that, if the polynomial has integer coefficients, then the values are integers as well; which in fact is a rather trivial statement. However, the opposite does not hold. That is, if the set of values of a polynomial consists of integers, this does not allow one to conclude that the polynomial itself has integer coefficients as well. To make the distinction, Pólya will present a first theorem. What is important to note here is that the opening paragraphs do not refer to any other mathematical (open) problem or theory. The question Pólya wants to deal with seems to be a problem on its own: one sees that a particular implication "If A then B" holds, and then wonders whether the inverse implication "If B then A" holds as well. If so, this yields an equivalence, a nice and potentially important result. As will become clear, however, the above appearance is deceptive. Pólya does not just tackle an isolated question here, but in fact *does* deal with existing and — at the moment of his writing — unsolved mathematical problems. He does not say so in the introduction. The reader, in other words, is kept in suspense. Why this problem as opposed to any other?

The theorem already announced says the following: given a polynomial $f(x)$ of degree $m$, if there are $m + 1$ values with absolute value smaller than $\frac{m!}{2^m}$, then the polynomial is not an integer polynomial. The proof relies on Lagrange's interpolation formula, which the reader is assumed to know, for it is immediately applied to the problem at hand. Although this might appear to be a detail, it does make an important difference. It is the distinction between the following two presentations:

[13] It can be found at the excellent website `http://www.digizeitschriften.de/`, where it is freely available for download or consultation. Unless otherwise indicated, the same holds for other papers mentioned in this section.

1. Taking into account Lagrange's interpolation formula, the following formula X holds.

2. Given Lagrange's interpolation formula Y, applied to the problem at hand, this transforms into the formula X.

Whereas in the latter case, the proof tries to be self-contained, in the former case, the scheme actually used by Pólya, it requires background knowledge. The rest of the proof is fairly straightforward, few steps remain to be inserted to obtain a more or less complete proof. Once again note that at this point in the paper the reader still has no idea where it is heading. After all, the first theorem does not solve the original problem, as we have not been told what happens when one (or more) of the $m+1$ values has an absolute value greater than $\frac{m!}{2^m}$.

It is common practice for a theorem to be followed by an illustration. Any good example will do, of course, and Pólya does not motivate his, although at a later stage his choice again turns out to have been well deliberated. The particular example is that of the following polynomial of degree 3: $f(x) = (x+1)x(x-2)+1$, and its four values 1, 1, $-1$, 1 for, resp, $x = -1, 0, 1, 2$. It is left to the reader to see why, which is immediate. Calculate $\frac{m!}{2^m}$ for $m = 3$, viz. $\frac{3}{4}$, and it is clear that $|-1| = |1| > \frac{3}{4}$, hence the condition of the proof is not satisfied. However, for polynomials of higher degree, it is impossible to have $m+1$ values 1 or $-1$, for, as Pólya observes (and no more explanation is given), in the ever growing series $\frac{1}{2}$, $\frac{1}{2}$, $\frac{3}{4}$, $\frac{3}{2}$, $\frac{15}{4}$, $\frac{45}{4}$, $\frac{315}{8}$, ..., $\frac{1}{2}\cdot\frac{2}{2}\cdot\frac{3}{2}\ldots\frac{m}{2}$, ... the fourth term is larger than 1.

It is worthwhile to look at this short passage in more detail. Firstly, it is not mentioned explicitly that the series presented is the result of calculating $\frac{m!}{2^m}$, for consecutive values of $m = 1, 2, 3, \ldots$ It has to be derived from the expression for the general term of the series, viz. $\frac{1}{2}\cdot\frac{2}{2}\cdot\frac{3}{2}\ \ldots\frac{m}{2}$. But why present it in this form? Why not simply $\frac{m!}{2^m}$? Secondly, reference is made to the fourth term, $\frac{3}{2}$, which is larger than one, so all $m+1$ values will be smaller than that value, and hence the condition of the proof is satisfied. Thirdly, it does not follow immediately that therefore the same holds for all values $m$ higher than 4. Now it does not really help to simply state that the numbers in the series are ever growing, unless the reader is prepared to check that $\frac{(m+1)!}{2^{m+1}} > \frac{m!}{2^m}$. That, however, becomes immediately obvious if one looks at the general term of the series: $\frac{1}{2}\cdot\frac{2}{2}\cdot\frac{3}{2}\ \ldots\frac{m}{2}$. Moving up one step, means multiplying with $\frac{m+1}{2}$, and if $m$ is larger than 4, then it is larger than 1, so the multiplication makes the number grow. With this, the proof is in a sense implicitly given, which explains why Pólya selected his particular way of representing the general term of the series.

The next theorem is stated without any context or explanation of its relevance or importance. In addition, a formerly unmentioned notion appears: irreducibility of polynomials. The theorem says that, if the integer polynomial $P(x)$ of degree $n$ has $n$ values, all different from 0 and in absolute value smaller than $\frac{(n-\frac{n}{2})!}{2^{(n-\frac{n}{2})}}$, then $P(x)$ is irreducible in integer polynomials. Similar to the proof of the first theorem, this proof too starts with a reference, in this case to Gauss's law, reading that if a polynomial has rational factors, then it must have integer factors. The steps that follow are again rather straightforward, and will not be dealt with in detail here. Should one wonder why it is not necessary in this theorem to talk about $n + 1$ values of a polynomial of degree $n$, as was the case in the first theorem, the answer is that one looks at the degree of a factor of the polynomial, say $k$, and that value will be less than $n$, hence there are $k + 1$ values that satisfy the condition stated in the theorem. The way Pólya formulates it is truly beautiful: he does not use $k$, as we have just done, but $m$, and without any further explanation writes down the following formula that $m$ must satisfy: $n - \frac{n}{2} \leq m < n$. It takes a moment's reflection to see that, if there is a factor of degree $m$, then the remaining factor has degree $nm$, so one or the other is larger than (roughly) half of $n$. A last curious element to note in the proof is its overall structure. It has the form of a *reductio*: let the polynomial be reducible, then it cannot be an integer polynomial, but the theorem says it is, hence it is irreducible.

Here too one gets the expected example illustrating the theorem. However, Pólya's choice seems rather peculiar: he considers polynomials of degree $n$ of the form $x(x-1)(x-2)\ldots(x-n+1)+\lambda n$. Why? Actually, this time we do get an answer, in a footnote that makes reference to a paper by P. Stäckel, "Arithmetische Eigenschaften ganzer Funktionen," learning that this example apparently was the very starting-point of Pólya's paper. This clearly suggests that Pólya intended his paper to begin exactly the way it did. He could just as well have done otherwise, starting with the problem, using some kind of canonised opening statement, like: "In this paper we will examine P. Stäckel's example in his paper so-and-so, and answer some open questions surrounding it." Furthermore, the background information given is really minimal. Pólya mentions Hilbert's irreducibility theorem — without quoting the theorem itself — and applies it directly to the example at hand, to state that there are integer values of $\lambda$ such that the polynomial is irreducible. Next he proposes to look at the special values 1 and $-1$ for $\lambda$. Surely one is immediately reminded of the previous example, where the values 1 and $-1$ were used as well. Parts of the paper that had before remained a bit "mysterious," now are situated and acquire their meaning, so to speak. But why these values? In the positive case, the polynomial becomes $x(x-1)(x-2)\ldots(x-n+1)+n$. To apply the second theorem we need $n$ values, all of them different from 0 and in absolute value smaller than $\frac{(n-\frac{n}{2})!}{2^{(n-\frac{n}{2})}}$. But that is now straightforward: take $x = 0, 1, \ldots, n-1$, then the first term disappears and the polynomial takes the value $n$, different from 0. So, if the condition $\frac{(n-\frac{n}{2})!}{2^{(n-\frac{n}{2})}} > n$ is satisfied, then we know that the polynomial is irreducible. None of this is mentioned in the paper, where one only reads this: "As soon as the condition is satisfied, then we can choose 1 or $-1$ for $\lambda$."

From the argumentative-rhetorical perspective, what follows is a true gem: for what values of $n$ is the condition indeed satisfied? We translate Pólya's words as faithfully as possible: "The condition will be satisfied for $n \geq 13$. Indeed, for $m \geq 7$, $\frac{m!}{2^m} > 2m > 2m-1$, $\frac{(m-1)!}{2^{m-1}} > 4$." What is going on here? The (again unmentioned) key is to see that $m$ plays the part of $n - \frac{n}{2}$, and so the condition indeed becomes $\frac{m!}{2^m} > 2m > 2m-1$, taking into account the even and the odd case for $n$. Take the first inequality, divide by $m$ on both sides, multiply by 2, and one finds $\frac{(m-1)!}{2^{m-1}} > 4$. If we read $m-1$ as the variable, then we know this series, and we now that is a steadily growing one. So we only need to check for the first value that is greater than 4. Do we need to make a lot of calculations? The answer is no, for remember the numbers $\frac{1}{2}$, $\frac{1}{2}$, $\frac{3}{4}$, $\frac{3}{2}$, $\frac{15}{4}$, $\frac{45}{4}$, $\frac{315}{8}$, ..., $\frac{1}{2}\cdot\frac{2}{2}\cdot\frac{3}{2}\ldots\frac{m}{2}$, ... The first value larger than 4 is $\frac{315}{8}$, the 6th term in the series, so $m-1 = 6$ or $m = 7$. And so $n \geq 13$.

Having covered almost two pages of the paper, we here end our partial analysis, for we believe we have made our point. Reading a mathematical text requires strategies for the reader to understand what is being said, what it is about, how arguments are presented, why proofs are considered important, in short, it takes an analysis that one typically associates with argumentation theory. The most important thing to note, however, is that our analysis takes as its subject the entire paper, not an isolated proof within it. To highlight the point we are trying to make here, it is worth taking a brief look at the work of Reviel Netz on the rhetoric of mathematics, which has been an important inspiration for the approach outlined here.

## 6. Netz and proofs as stories

The Stanford historian of science Reviel Netz some time ago launched the intriguing idea of reading and treating mathematical texts, including proofs, as stories.[14] We have found that this proposal shares with ours the fundamental thought that an argumentative analysis should not merely apply to individual proofs, but also to coherent collections of proofs. A prime example is the thirteen books of Euclid's *Elements*. To avoid misunderstanding, the object is not to reduce mathematics to literary prose, but rather to draw attention to the rhetorical elements inherent to mathematical texts. A short comment.[15]

An individual proof can always be assigned a tree-like structure, showing its skeleton. Picture the conclusion as the "root" of the tree, the premises as the extremes of "branches," and a story-line becomes readily visible. A concrete example. Suppose we want to show that the sum of the first $n$ natural numbers equals $\frac{n.(n+1)}{2}$. A standard proof, based on mathematical induction, proceeds as follows:

(a) Basis: suppose that $n = 0$, then

- (a1) $1 + 2 + 3 + \ldots + n = 0$,[16] and
- (a2) $\frac{n.(n+1)}{2} = 0$, and
- (a3) so they are equal.

(b) Induction step: suppose the statement holds for $n$, so $1 + 2 + 3 + \ldots + n = \frac{n.(n+1)}{2}$

(c) Add to both sides $n + 1$: $1 + 2 + 3 + \ldots + n + (n + 1) = \frac{n.(n+1)}{2} + (n + 1)$

(d) The right-hand side can be transformed into:

- (d1) $\frac{n.(n+1)}{2} + (n + 1) = (n + 1).(\frac{n}{2} + 1)$
- (d2) $\frac{n.(n+1)}{2} + (n + 1) = \frac{(n+1).(n+2)}{2}$
- (d3) $\frac{n.(n+1)}{2} + (n + 1) = \frac{(n+1).((n+1)+1)}{2}$

(e) Put this together, and one finds: $1+2+3+\ldots+n+(n+1) = \frac{(n+1).((n+1)+1)}{2}$, which is precisely the statement to be proven for $n + 1$.

[14] See, e.g., Netz (1999). However, Netz has definitely not been the only one to deal with the topic. We mention Thomas (2000) and (2002), where the comparison between stories and proofs is examined in depth, as well as Otte & Panza (1997) and Rotman (2000), for a semiotically inspired approach.

[15] For a more elaborate treatment, see Van Bendegem ([to appear I]).

[16] Although something like $\sum_{n=0}^{0} n = 0$ would indeed appear less paradoxical, do note that exactly the same is expressed here. Moreover there is no such thing as the first zero naturals, which renders this formula utterly meaningless (just like 0!).

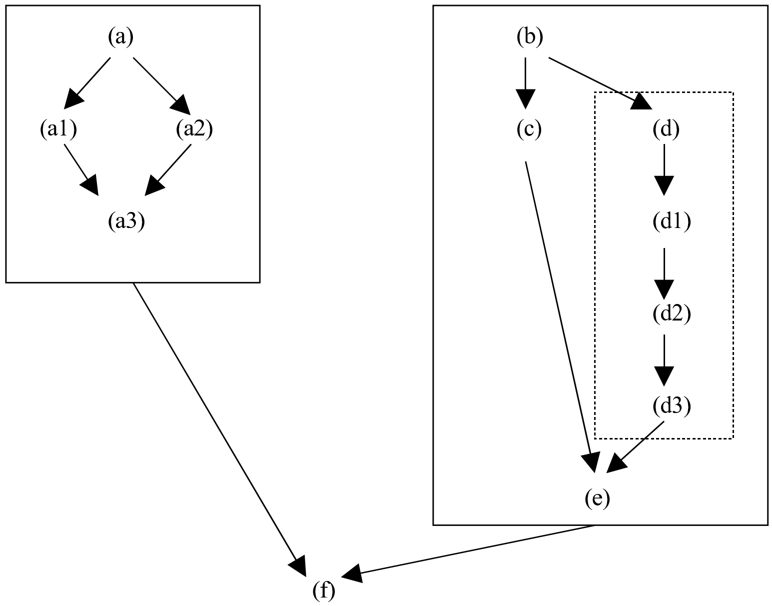


*Figure 2.* Tree structure of a proof for $\frac{n.(n+1)}{2}$

(f) By mathematical induction, for all $n$, $1+2+3+\ldots+n = \frac{n.(n+1)}{2}$.

The tree structure of this proof corresponds to that of figure 2.

We can, of course, discuss the identification of the correct elements in the tree, but the idea itself is clear enough. If one is willing to show some generosity in interpretation, the four steps in the small rectangle, viz. (d) up to (d3), can be seen as a "subplot" of the story. But even without such generosity, it is clear that the steps from (d) to (d3) do exhibit a degree of independence from the overall proof. Or, to reformulate this observation in mathematical terms, it indicates that these steps could very well have been grouped into a separate proof, in other words, as a lemma that preceeds the actual proof. And what holds for an individual proof, *a fortiori* holds for a set of interconnected proofs, say, a mathematical treatise, handbook or survey. All individual proofs have their own structure, but theorems refer to one another, mostly in non-linear structures. This opens up interesting possibilities as to the organisation of the theorems and proofs. A concrete example: imagine a situation where the proof of a theorem $S_1$ requires a set of preparatory proofs $A_1, A_2, A_3, \ldots, A_n$, while for a theorem $S_2$ the proofs $B_1, B_2, B_3, \ldots, B_m$ are needed. A possible 'storyline' is a linear one: present all $A_i$'s, followed by $S_1$, and then all $B_i$'s, followed by $S_2$. However, it could very well be that, say, $A_1, A_2, B_1$ and $B_3$ are themselves connected because, e.g., they are all about arithmetical properties, such that they could be brought together in a separate introductory chapter "Arithmetical Properties." We would like to emphasize that any choices of this kind can have long-ranging effects. By grouping results in this (rather than another) way, it is suggested that the topics treated therein have a certain independence from other topics treated. In some cases, this kind of grouping of theorems under a separate heading can give rise to the creation of a separate mathematical domain — because people see connections where perhaps little or none were seen before — which might have an impact on mathematical development as a whole.

As an example, take the origin of the theory of complex numbers. In the first instance, the introduction of imaginary numbers was confined to the treatment of solutions of third degree polynomial equations. While denying its importance, their creator, Cardano, had already indicated that imaginaries could be studied on their own (Van Bendegem [to appear II]). That is precisely what Bombelli later did, demonstrating that indeed this does form a topic on its own. When eventually the link was made with the number of solutions of an equation and its degree, it became much more than a topic on its own, it became a basic constituent element of mathematics. Compare, as a matter of contrast, such beautiful formulas as $(x+y+z)^7 - (x^7+y^7+z^7) = 7(x+y)(x+z)(y+z)[(x^2+y^2+z^2+xy+xz+yz)^2+xyz(x+y+z)]$, that are extremely useful in solving specific problems like Fermat's Last Theorem for the special case $n = 7$ (Lamé 1839, and simplified by Lebesgue), without ever having become a topic of study on their own. As a remaining challenge, it would be an excellent argument in support of our thesis to rewrite a set of (historical) papers in such a way that other concepts "catch the mathematician's eye," thereby potentially giving rise to alternative developments.

## 7. Conclusion

We have tried to establish that in order to clarify the connections between proofs and arguments, an analysis of individual proofs in terms of a single type of argument, as roughly carried through in Toulmin-style exercises, does not suffice. What needs to be supplemented, is mathematical context. For except in very poor contexts (e.g., that of mathematical puzzles), mathematical arguments as such, i.e. in isolation from other mathematical arguments, are nonexistent. Instead, the natural habitat of most individual arguments are trains of arguments, adding up to interconnected theorems, theories and eventually entire fields. We have tried to illustrate this fundamental point by presenting a number of cases, concluding that, if one even wants to begin to do justice to full mathematical practice, large-scale organization of research programmes (section 4) as well as (alternative) modes of presentation pursuing low-level research goals (section 5) are to be actively addressed. This, in its turn, will ever more prominently bring into focus the informal, and indeed rhetorical aspects attached to arguing in mathematics (section 6), giving currency to the humanist case for mathematics as an open-ended endeavour.

## Authors' Vitae

*Jean Paul Van Bendegem*
(°1953) and

*Bart Van Kerkhove*
(°1972) are director and research assistant respectively at the Centre for Logic and Philsophy of Science of Vrije Universiteit Brussel, Belgium. Their research is mainly focused on the philosophy of mathematical practice, a topic on which they have published several (joint) papers, most recently "Pi on Earth, or Mathematics in the Real World", *Erkenntnis* 68(3):421-35 (2008), as well as an edited volume, *Perspectives on Mathematical Practices*, Springer (2007). Further research interests include (strict) finitism, philosophical logic, semiotics, and the socio-history of science.

Website: `http://www.vub.ac.be/CLWF`

E-mail: `jpvbende@vub.ac.be`, `bvkerkho@vub.ac.be`.

## Acknowledgements

Our appreciation goes to three anonymous referees, whose profound remarks on an earlier version helped us to improve the paper considerably.

## References

Aberdein, A. Managing Informal Mathematical Knowledge: Techniques from Informal Logic. *Lecture Notes in Artificial Intelligence*, 4108:208–21, 2006.

Aberdein, A. The Informal Logic of Mathematical Proof. In B. Van Kerkhove and J. P. Van Bendegem, eds., *Perspectives on Mathematical Practices*, pages 135–51, Springer, Dordrecht, 2007.

Aigner, M. and G. M. Ziegler. *Proofs from The Book*. Springer, New York, 1998.

Aschbacher, M. The Status of the Classification of the Finite Simple Groups. *Notices of the AMS*, 51(7):736–40, 2004.

Cipra, B. *What's Happening in the Mathematical Sciences*. Volume 3. American Mathematical Society, Providence, RI, 1996.

Davis, P. J. Fidelity in Mathematical Discourse: Is One and One Really Two?. *American Mathematical Monthly* 79(3):252–63, 1972.

Kline, M. *Mathematical Thought from Ancient to Modern Times*. Oxford University Press, New York, 1990. (Three volumes, second printing of the 1972 original edition.)

Kolodziejczyk, L. and N. Thapen. The Polynomial and Linear Hierarchies in Models Where the Weak Pigeonhole Principle Fails. *The Journal of Symbolic Logic* 73(2):578–92, 2008.

Lamé, G. Mémoire sur le Dernier Théorème de Fermat. *Comptes Rendus de l'Académie des Sciences* 9:45–46, 1839.

Nelsen, R. B. *Proofs Without Words. Exercises in Visual Thinking*. The Mathematical Association of America, Washington, DC, 1993.

Nelsen, R. B. *Proofs Without Words II. More Exercises in Visual Thinking*. The Mathematical Association of America, Washington, DC, 2000.

Netz, R. *The Shaping of Deduction in Greek Mathematics. A Study in Cognitive History*. Cambridge University Press, Cambridge, 1999.

Otte, M. and M. Panza. *Analysis and Synthesis in Mathematics: History and Philosophy*. Kluwer Academic Publishers, Dordrecht, 1997.

Pólya, G. Verschiedene Bemerkungen zur Zahlentheorie. *Jahresbericht der deutschen Math.-Vereinigung* 28:31–40, 1919.

Pólya, G. *Mathematics and Plausible Reasoning. Volume I: Induction and Analogy in Mathematics*. Princeton University Press, Princeton, NJ, 1973. (eighth printing of the 1954 original)

Queneau, R. *Bords. Mathématiciens. Précurseurs. Encyclopédistes*. Hermann, Paris, 1978. (first edition 1963)

Rotman, B. *Mathematics as Sign. Writing, Imagining, Counting*. Stanford University Press, Stanford, 2000.

Solomon, R. On Finite Simple Groups and Their Classification. *Notices of the AMS*, 42(2):231–9, 1995.

Solomon, R. A Brief History of the Classification of the Finite Simple Groups. *Bulletin (New Series) of the AMS*, 38(3):315–52, 2001.

Thomas, R. Mathematics and Fiction I: Identification. *Logique et Analyse*, 43:301–40, 2000.

Thomas, R. Mathematics and Fiction II: Analogy. *Logique et Analyse*, 45:185–228, 2002.

Thurston, W. P. On Proof and Progress in Mathematics. *Bulletin (New Series) of the AMS*, 30(2):161–77, 1994.

Van Bendegem, J. P. Elements for a Rhetoric of Mathematics: How Proofs Can Be Convincing. To appear in C. Dègremont, L. Keiff and H. Rückert, eds., *Festschrift in Honour of Shahid Rahman*, College Publications, London.

Van Bendegem, J. P. 'What-if' Stories in Mathematics: An Alternative Route to Complex Numbers. To appear in M. Khissine et al., eds., *Liber Amicorum Marc Dominicy*.

Van Kerkhove, B. and J. P. Van Bendegem. Pi on Earth, or Mathematics in the Real World. *Erkenntnis*, 68(3):421-435, 2008.