International Newsletter on the Teaching and Learning of Mathematical Proof

La Lettre de la Preuve
Printemps 2003

On line material

OBSERVATION AND DESIGN IN MATHEMATICAL PROOFS

Willi Dörfler

Institut für Mathematik, Universität Klagenfurt, Austria

La lettre de la Preuve : http://www-didactique.imag.fr/preuve

Observation and design in mathematical proofs

Willi Dörfler

Institut für Mathematik, Universität Klagenfurt, Austria

Introduction

This contribution is located in the context of the philosophy of mathematics by the American philosopher and pragmatist Ch. S. Peirce. Yet, it is readable and understandable without a detailed knowledge of the stance taken by Peirce. The interested reader might consult the papers Dörfler (2003a, 2003b) or Hoffmann (2001, 2002). This especially holds for the notion of diagram and diagrammatic reasoning which were introduced by Peirce to explain, on the one hand, the stringency of mathematical proofs and, on the other hand, the possibility of inventions and constructions in mathematics, or what he calls "surprising observations". Thus he says (in Peirce, Collected Papers 3.363):

> It has long been a puzzle how it could be that, on the one hand, mathematics is purely deductive in its nature, and draws its conclusions apodictically, while on the other hand, it presents as rich and apparently unending a series of surprising discoveries as any observational science. Various have been the attempts to solve the paradox by breaking down one or other of these assertions, but without success. The truth, however, appears to be that all deductive reasoning, even simple syllogism, involves an element of observation; namely, deduction consists in constructing an icon or diagram the relations of whose parts shall present a complete analogy with those of the parts of the object of reasoning, of experimenting upon this image in the imagination, and of observing the result so as to discover unnoticed and hidden relations among the parts. ... As for algebra, the very idea of the art is that it presents formulae, which can be manipulated and that by observing the effects of such manipulation we find properties not to be otherwise discerned. In such manipulation, we are guided by previous discoveries, which are embodied in general formulae. These are patterns, which we have the right to imitate in our procedure, and are the icons par excellence of algebra.

From this it already comes clear that a diagram might be of a great variety: geometric figures and algebraic expressions as well. For short, diagrams in Peirce are special (iconic) signs which have a clearly defined and recognizable structure and which can be manipulated according to (conventional) rules for transformations and compositions, cf. again the above mentioned papers. The crux of all that is that empirical and perceptive observation becomes a decisive part of mathematical reasoning, of devising and understanding proofs and mathematical arguments. Mathematical reasoning in this view is not so much the handling of abstract ideas in one's mind but the observation of the effects of one's manipulations of diagrams. The mathematical ideas rather reside in the invention of diagrams and of their fruitful manipulations, transformations, compositions. Mathematics in this sense studies the general properties and regularities of certain diagrams and of the operations with them. In accordance with the triadic sign concept of Peirce those diagrams will be interpreted by their users in many different ways and will be related to "objects" (in the sense of Peirce) also in various ways. In a way, I am analysing here only the sign aspect (representamen) of that Peircean triad "sign, object, interpretant" but apparently this for mathematics is of crucial importance.

From diagrammatic reasoning derives also the absolute reliability and security of mathematics, its so-called logical necessity. This differentiates observation of diagrams also from empirical observation in the natural sciences. Diagrammatic observation "sees" that a certain relationship will hold in all conceivable instances of the respective type of diagrams. This is enabled by the generic character of the (mathematical) diagrams: each single instance or token fully presents the respective type according to an adequate perspective on the token. Like, say, any inscription of the letter "a" presents that letter (as a type of inscriptions under a certain perspective). Finally, it should be emphasized that diagrammatic reasoning is very much different from algorithmic calculations. Though it is rule based it needs creativity and inventiveness like composing music.

Observing Diagrams

In this section I will present some examples which hopefully offer to the reader the experience that mathematical proofs in many cases depend on the observation of structural relationships and regularities within transformations of diagrams. Other examples can be found in Dörfler (2003a). In all examples the results of previous "experiments" with diagrams are used as established formulae or "theorems".

There is the surprising result of: 11x11=121, 111x111=12321, 1111x 1111=1234321 etc. This can be "explained" by observing a diagram like:

1	1	1	1	1	1	1	x	1	1	1	1	1	1	1
1	1	1	1	1	1	1								
	1	1	1	1	1	1	1							
		1	1	1	1	1	1	1						
			1	1	1	1	1	1	1					
				1	1	1	1	1	1	1				
					1	1	1	1	1	1	1			
						1	1	1	1	1	1	1		
1	2	3	4	5	6	7	6	5	4	3	2	1		

One of the rules used here is the decimal multiplication algorithm which in itself does not predict the observed relationships in the above diagrams. The "understanding" of the surprising results derives from recognizing the pattern of 1's which is produced by the algorithm. The usual common interpretation of the symbols might be helpful but the essential point consists in the perceptive observation of the outcome of one's operations on the diagrams. These would hold even if there were no interpretation of the symbols as numbers. A precondition for this diagrammatic reasoning clearly will be a close familiarity with the diagrams and proficiency in their operations. This possibly sheds new light on the role of "calculations" conceived in a wider sense as intelligent and creative operations with diagrams. Based on these first observations there is a rich space of further diagrammatic experiments and thought experiments with those diagrams. There is also the possibility of changing the diagrammatic rules, e.g. by choosing different bases for the place value system.

Also the next example – as the others as well – is well known and only serves the purpose of orienting the attention of the reader to the role of perception, observation, pattern recognition and manipulation of concrete inscriptions as a constitutive part of mathematical thinking.

The young Gauss is reported to have found the sum of the first 100 positive integers by thinking of those numbers as being written down in the following way

1	2	3	4	…	49	50
100	99	98	97	…	52	51

and adding the two numbers in each of the 50 columns to get $50 \times 101 = 5050$ as the required sum. This is very similar to our first example: a certain recognized pattern in a diagram gives the result. Here the generic character (for even numbers) can be seen: a thought experiment with the respective diagram gives the formula $((n/2)x(n+1))$. Further experiments with those diagrams will lead to another more general diagram for arbitrary n, like the following:

1	2	3	4	5	6	7
7	6	5	4	3	2	1

I do not deny that an understanding of the involved symbols as natural numbers is helpful or even necessary for recognizing the relevant pattern. But for the latter a certain regularity, namely constant sum in the columns, is most important, and that is not inherently related to natural numbers. Thus, the diagram is added to the known properties of natural numbers and enlarges the knowledge about them. In a similar way one can analyze many other number patterns like triangular, square, rectangular numbers. In all cases besides symbolic presentations graphic ones using arrays of dots is another kind of diagrammatic reasoning based on experiments with and observation of diagrammatic structures. To that already point names like "triangular", "square" or "rectangular".

Within Linear Algebra there is a wealth of examples for diagrammatic reasoning. The basic diagrams there are matrices and their operations. Consider $A=(a_{ij})$ an $(m\times n)$ - matrix and $\boldsymbol{a}=(a_j)$ an $(n\times 1)$ - matrix (vector). Then the i - th component of the product $A\boldsymbol{a}$ is

$$a_{i1}a_1 \quad + \quad a_{i2}a_2 \quad + \quad \cdots \quad + \quad a_{in}a_n$$

or more detailed the vector $A\boldsymbol{a}=(b_i)$ is given as:

$$\begin{array}{ccccccccc} b_1 & = & a_{11}a_1 & + & a_{12}a_2 & + & \cdots & + & a_{1n}a_n \\ b_2 & = & a_{21}a_1 & + & a_{22}a_2 & + & \cdots & + & a_{2n}a_n \\ & & & & & \cdots & & & \\ b_m & = & a_{m1}a_1 & + & a_{m2}a_2 & + & \cdots & + & a_{mn}a_n \end{array}$$

An empirical investigation of this diagram exhibits a column-wise regularity which can be expressed as

$$A\boldsymbol{a} = a_1\boldsymbol{a}_1 + a_2\boldsymbol{a}_2 + \cdots + a_n\boldsymbol{a}_n$$

where $\boldsymbol{a}_j=(a_{ij})$ is the j - th column-vector of A. This is a result which is a stringent consequence of the operation rules for matrices and the diagram above cannot be doubted, it is an apodictic argument though (or possibly because of) being based on "pattern recognition".

Once such a "pattern" is established as a formula or theorem it can fruitfully be used to derive further consequences. Assuming $\boldsymbol{a}$ to be the i - th unit vector $e_i\,(a_j=0 \text{ for } j\neq i, a_i=1), i=1,\ldots,n$, leads to $Ae_i=\boldsymbol{a}_i$ which of course can be recognized from other diagrams also. Here it becomes even more prominent that the important thing are the operational rules and not so much the (referential) meaning of the symbols manipulated. We only use our knowledge how to operate with the symbols. But still it is not a meaningless, purely formalistic game: we discover surprising and fascinating relationships for the diagrams. Thus diagrams play here manifold roles. They are, on the one hand, the objects of reasoning properties of which are detected and described (by new diagrams). On the other hand, diagrams are the means for mathematical reasoning by which relationships and regularities become observable patterns.

As another example we study one of the proofs of Cramer's rule for the solution of a regular square system of linear equations $Ax = b; A = (a_{ij})$ an nxn matrix, $x = (x_i)$ the solution vector, $b = (b_i)$ the right-side vector. Then by assumption the inverse A^{-1} (with $AA^{-1} = A^{-1}A =$ identity matrix) exists and from previous diagrammatic operations one knows that $A^{-1} = (A_{ji} / |A|)$ where $|A|$ is the determinant, and A_{ji} is the cofactor of a_{ji} in A. Then $x = A^{-1}b$ and therefore

$$x_i = (1/|A|)(A_{1i}b_1 + A_{2i}b_2 + \ldots + A_{ni}b_n)$$

Now $A_{1i}b_1 + A_{2i}b_2 + \ldots + A_{ni}b_n$ is observed to be the result of expanding the determinant of the following matrix A_i by the i - th column

$$A_i = \begin{pmatrix} a_{11} & \ldots & a_{1,i-1} & b_1 & a_{1,i+1} & \ldots & a_{1n} \\ a_{21} & \ldots & a_{2,i-1} & b_2 & a_{2,i+1} & \ldots & a_{2n} \\ & \ldots & & & & & \\ a_{n1} & \ldots & a_{n,i-1} & b_n & a_{n,i+1} & \ldots & a_{nn} \end{pmatrix}$$

since A_{ji} is just the appropriate cofactor resulting from deleting the j - th line and i - th column in A_i or equivalently in A. Thus $x_i = |A_i| / |A|$. This clearly is recognized by observing invariant patterns when carrying out diagrammatic operations or experiments. For the latter, intimate experience with those diagrams and their previously observed properties is indispensable. Of course, several experimental trials with the diagrams will be necessary before a useful pattern will be discovered. In any case it is scrutinizing the diagrams which is at the core of "inventing" the proof. In the hindsight, this might then be presented as the "idea of the proof". We should therefore not expect our students at any level to be able to independently produce proofs without preceding intensive work on the respective diagrams. The reader might interpret this for instance in the case of (Euclidean) geometric proofs. The reader is also encouraged to have a look in a standard text-book on Linear Algebra and to read some of the proofs under the pretext of diagrammatic reasoning. He/she will observe again and again the importance of observing and recognizing patterns of relationships in the produced diagrams which are constitutive for the respective proof. Instructive examples are: row rank equals column rank; matrix of a linear transformation; basis change for linear transformations. But of course already the basic properties of the matrix operations are good examples for diagrammatic reasoning.

Reading a finished (diagrammatic) proof demands first of all proficiency in recognizing patterns in diagrams. Devising a proof mostly is based on inventing new diagrams or parts of them. This becomes most clear in geometric proofs in the form of auxiliary lines and figures. Here I will refrain from studying geometric proofs because the diagrammaticity of mathematical reasoning might be more unexpected in other fields. For calculus see Dörfler (2003a).

As another example for a crucial invention I take the standard proof of the Cauchy-Schwarz-Inequality for an inner-product (a,b), i.e. $(a,b)^2 \leq (a,a)$ (b,b). One invents a new diagram $(a+xb, a+xb)$, x real number, and then observes the transformation: $0 \leq (a+xb, a+xb) = (a,a) + 2x(a,b) + x^2(b,b)$ which is using the conventional properties of the inner-product. From diagrammatic reasoning with quadratic polynomials one now knows that $b^2 - 4ac \leq 0$ if $ax^2 + bx + c \geq 0$ for all x. And this gives for the above diagram

$$4(a,b)^2 - 4(a,a)(b,b) \leq 0$$

which is the desired inequality.

Clearly, this kind of diagrammatic reasoning presupposes intimate acquaintance with the handling of symbols and with ascribing generality to the respective expressions. But still the diagrammatic operations and their observation adds to all this and constitutes the core of the proof, its stringency and security. Thus, mathematics cannot be reduced to diagrammatic reasoning but the latter is an essential component of its specific quality and character. Specifically, having at hand a great inventory of diagrams, diagrammatic relationships and operations is a precondition for mathematical inventiveness and productive ideas. The latter very often are rich and productive diagrams of some sort. Take as an example the Pascal triangle in Combinatorics or possibly simple number relations in the context of developing number sense.

Design

In this section I will present examples for a specific type of proof. It is those proofs which consist in the purposeful design or construction of a certain kind of diagrams or in the proof of the possibility of such a construction. In a sense, those are constructive existence proofs by exhibiting diagrams with the desired property or properties. A simple example is the proof that between any two fractions m/n and p/q there is another one: Assuming $m/n < p/q$ we find $mq < np$ because of $mq/nq < np/nq$. Then for any k between $2mq$ and $2np$ the fraction $k/2nq$ will be a required fraction. Or, the Euclidean proof that for any given set of prime numbers we can find one not in this set is also of that kind.

The next example on the first glance does not give the impression that in essence it is the design of diagrams which is at the centre of the proof of the theorem. It is the well-known theorem by Kronecker about the existence of roots for polynomials. More technically the theorem reads as follows. For any polynomial $P(x)$ over a field F (i.e. the coefficients of P are elements of F) there is an extension-field F_1 of F where P has a root (i.e. in F_1 there is an element r with $P(r)=0$ over F_1). Thereby one can assume additionally that P is irreducible over F (i.e. P is not the product of two polynomials over F each of degree 1 at least). The proof starts by considering the ring $F(x)$ of all polynomials over F which can be considered to be a class of diagrams in the sense used here. Then the general construction of the field $F(x)/P(x)$, of $F(x)$ modulo $P(x)$, is employed which can be introduced as consisting of all equivalence classes of $F(x)$ modulo $P(x)$. Thereby $p_1 = p_2(P)$ if $p_1 - p_2$ is a multiple of P in $F(x)$. Denoting by $[p]$ the class of $p \in F(x)$ the field operations on the classes are given by $[p_1]+[p_2]=[p_1+p_2]$ and $[p_1]x[p_2]=[p_1xp_2]$. The latter definitions have diagrammatic character but the notion of an equivalence class itself is not of a diagrammatic quality. Yet, the whole "construction" can be described easily in a diagrammatic view. In each class of $F(x)/P(x)$ there is a unique polynomial p of degree less than the degree of P. If n is the degree of P then $F(x)/P(x)$ can be viewed as the set of all polynomials $a_0 + a_1x + \dots a_{n-1}x^{n-1}$ over F with the usual addition and a certain multiplication. The latter results from p_1xp_2 (polynomial product) by reduction modulo P, i.e. it is the remainder of the division of p_1xp_2 by P. By various diagrammatic manipulations one demonstrates that those operations for those diagrams satisfy all the properties of a field. The field F clearly is contained in the new field $F_1 = F(x)/P(x)$ and thus P can be viewed as being a polynomial over F_1. Among all the diagrams of F_1 there is the special diagram x (i.e. we have $a_0 = a_2 = \dots = a_{n-1} = 0$ and $a_1 = 1$), and for this diagram we find according to the diagrammatic rules of F_1 that $P(x)=0$ in F_1 since $P(x)=1 \cdot P(x)+0$, i.e. 0 (the zero polynomial in F_1) is the remainder when dividing P by P. But this is just the same as saying that x is a root of P in F_1. To summarize: the proof can be interpreted in a diagrammatic way as the design of a class F_1 of diagrams containing the elements of F for which a sum and a product can be defined such that F_1 is an extension field of F; and in F_1 there is a diagram $r(=x)$ which is a root of P over F_1. The important property of this proof by design is that we can construct a diagram r which is a root of P (this is easy: just say that r has the property $P(r)=0$) and which is element of an extension field (this is the hard and possibly surprising part). Ontologically, the theorem and its proof are not about abstract objects but about perceivable, observable and materially manipulable objects, viz, the diagrams $a_0 + a_1x + \dots + a_{n-1}x^{n-1}$.

The best known special case of the above of course are the complex numbers where $F = R$ (the real numbers) and $P(x) = x^2 + 1$. Thus the resulting diagrams are of the form $a + bx$ and $x = i$ is a root of $x^2 + 1$ in $F_1 = C$. The product in F_1 results from $(a + bx)(c + dx) = ac + (ad + bc)x + bdx^2 = (ac - bd) + (ad + bc)x + bd(x^2 + 1)$ which in F_1, i.e. modulo $x^2 + 1$, is $(ac - bd) + (ad + bc)x$. The reader will recognize the usual product in C where we write $a + bi$ instead of $a + bx$. The diagrams in C can be designed more directly, of course, without the use of the polynomials. This proceeds by considering all diagrams of the form $a + bi$, by defining a sum and a product for them based on $i^2 = -1$ (a stipulated diagram again) and by demonstrating via diagrammatic manipulations that thereby results a field. Focusing on the diagrams, their design and their operations instead of looking for "numbers" which are denoted by those diagrams turns this construction into a rational and even perceivable and observable one. The complex numbers thereby loose their common imaginary and mythical quality. Thus the diagrammatic point of view contributes to demystifying mathematics. Of course, there remains the infinity of R which is beyond diagrammatic means. Yet, on the level of C this does not pose specific problems.

To make the design of a root of $P(x)$ and of a field containing it even more transparent I choose the specific case of $F = Z_5$, i.e. the field of residue classes of Z modulo 5 which we denote for the sake of simplicity of writing by 0,1,2,3,4. Consider the polynomial $P(x) = x^2 + 2$ which easily is seen to have no root in Z_5 since the squares in Z_5 are 0,1 and 4 ($x^2 + 1$ would have 2 as a root since $4 + 1 = 0$ in Z_5). The elements of $F(x)/P(x)$ are therefore the diagrams $a + bx, a$ and b in Z_5, which are 25 elements among them all of Z_5 and, for example, $2x, 3x, 4x, 2 + 3x$, etc. For the sum, we have for example $(2 + 3x) + (3 + x) = 0 + 4x = 4x$; and for the product $(2 + 3x)(3 + x) = 1 + 2x + 4x + 3x^2 = 1 + x + 3x^2 = x + 3(x^2 + 2) = x$ modulo P. The latter more easily is obtained by using $x^2 = -2 = 3$ in Z_5 or better in F_1. It is then a matter of diagrammatic reasoning to convince oneself that those newly designed diagrams with their operations of sum and product have all the properties of a field. Most of them are direct consequences of the respective properties holding in Z_5. For the multiplicative inverse one has to solve the equation $(a + bx)(c + dx) = 1$ with a,b given for $c,d \in Z_5$. If $b = 0$ then $c = 1/a$ and $d = 0$; otherwise $c = a/(a^2 + 2b^2)$ and $d = (-b)/(a^2 + 2b^2)$ (observe that $a^2 + 2b^2 \neq 0$ for all $a,b \in Z_5$ not both zero). In this (finite) case one has a complete survey of all diagrams and there is absolutely no need for abstract objects which the diagrams possibly stand for. At least in these cases the mathematics is about the writing and manipulating of diagrams according to conventional rules which derive from specific purposes and intentions which can be viewed to be a possible interpretant of the diagrams (the signs) in the sense of Peirce. Possibly one has then to take the diagrams as their own objects to complete the triadic sign relationship of Peirce.

A similar analysis could be carried out for many other mathematical "constructions". I just mention some more examples: direct products of algebraic structures (design is the writing of ordered pairs); design of finite geometries; existence of (combinatorial) graphs with certain properties.

Conclusion

I hope the reader has got an idea of what is meant by diagrammatic reasoning and of its power and usefulness in mathematics. But I hasten to emphasize that mathematics cannot and should not be reduced to diagrammaticity. There are powerful ways of mathematical thinking and reasoning which appear to evade diagrammatic methods, see Dörfler (2003b). Of very great interest also for the learning of mathematics possibly is the intricate interplay of diagrammatic and other ways of presenting mathematical ideas, their relationships and differences.

References

Dörfler, W. (2003a). Diagrams as Means and Objects of Mathematical Reasoning. To appear in: Developments in Mathematics Education in German-speaking Countries. Selected Papers from the Annual Conference on Didactics of Mathematics.

Dörfler, W. (2003b). Diagrammatic Thinking: Affordances and Constraints. To appear in: Activity and Sign – Grounding Mathematics Education (eds. M. Hoffmann, J. Lenhard and F. Seger).

Hoffmann, M. H. G. (2001). Skizze einer semiotischen Theorie des Lernens. In: Journal für Mathematik-Didaktik 22(3/4): 231-251.

Hoffmann, M. H. G. (2002). Peirce's "Diagrammatic Reasoning" as a Solution of the Learning Paradox. In: The Quiet Revolution: Essays on Process Pragmatism. Ed. Guy Debrock. Amsterdam et al.: Rodopi Press: 147-174.

Peirce, CH. S. (1931-1958): Collected Papers, vol. I-VIII. Harvard University Press, Cambridge, Mass.

The papers Dörfler (2003a,b) will be sent by e-mail on request.

Willi Dörfler
Institut für Mathematik
Universität Klagenfurt
A-9020 Klagenfurt, Austria

Reactions? Remarks?

The reactions to the contribution of Willi Dörfler

will be published in the Summer 2003 Proof Newsletter