The Review of Symbolic Logic
Volume 10, Number 1, March 2017

'CHASING' THE DIAGRAM—THE USE OF VISUALIZATIONS IN ALGEBRAIC REASONING

SILVIA DE TOFFOLI
Department of Philosophy, Stanford University

Abstract. The aim of this article is to investigate the roles of commutative diagrams (CDs) in a specific mathematical domain, and to unveil the reasons underlying their effectiveness as a mathematical notation; this will be done through a case study. It will be shown that CDs do not depict spatial relations, but represent mathematical structures. CDs will be interpreted as a hybrid notation that goes beyond the traditional bipartition of mathematical representations into diagrammatic and linguistic. It will be argued that one of the reasons why CDs form a good notation is that they are highly mathematically tractable: experts can obtain valid results by 'calculating' with CDs. These calculations, take the form of 'diagram chases'. In order to draw inferences, experts move algebraic elements around the diagrams. It will be argued that these diagrams are dynamic. It is thanks to their dynamicity that CDs can externalize the relevant reasoning and allow experts to draw conclusions directly by manipulating them. Lastly, it will be shown that CDs play essential roles in the context of proof as well as in other phases of the mathematical enterprise, such as discovery and conjecture formation.

§1. Introduction. Visual representations of various natures are ubiquitous in mathematics, as well as in many other human activities. In the literature, these representations have been commonly labeled 'visualizations'. This term is actually vague: It has been employed in many different contexts in order to refer to various processes. As Carter (2010, p. 3) points out, the two main ways of using the term 'visualization' are to refer to "mental pictures" (or, more generally, "mental models") and to material representations. In the present study, the second meaning of the term will be endorsed and thus visualization will be considered as an 'externalization' of thought. Visualization can therefore be analyzed in order to shed light on specific aspects of human reasoning. The underlying assumption is that reasoning is *heterogeneous*, that is, not solely constituted by linguistic content. As Shin sums up (2004, p. 92), "all of us engage in and make use of valid reasoning, and in the process of reasoning human beings obtain information through many different kinds of media, including diagrams, maps, smells, sounds, as well as written or spoken statements."

In this article, I propose a way of interpreting some specific cognitive and epistemic artefacts pertaining to visualization: commutative diagrams (CDs) in homological algebra. These diagrams are highly formalized diagrams used in algebra. They do not depict geometric objects, but they represent algebraic relations. Still, even if the pictorial element is absent, CDs exhibit essential spatial properties that make certain relations of the represented algebraic structure visible. CDs form an extremely perspicuous mathematical notation that allows mathematicians to perform specific epistemic operations, that is, calculations aimed at the production of mathematical results. I will describe CDs as a hybrid mathematical notation and unveil the reasons underlying their effectiveness. To do so,

Received: December 21, 2015.

doi:10.1017/S1755020316000277

I will start by developing a general methodology to evaluate mathematical notations, and then I will apply it to the case of CDs. Therefore, the main contribution of this article is twofold: (i) a theoretical explanation of the effectiveness of CDs in mathematics[1] and (ii) a methodological contribution to evaluate mathematical notations in general.

More specifically, through a case study I will argue for the following five main claims: (1) CDs form an effective notational system.[2] In order to sustain this claim, I will analyze their uses in specific mathematical contexts. I will apply to their analysis a general methodology to evaluate mathematical notations, which I will develop. (2) CDs function like geographic *maps* of abstract spaces: As I will describe in §4.2, they describe an algebraic ‘landscape’. Moreover, their use consists in a constant feedback between diagrams and text. (3) CDs are ‘dynamic’, that is, their functionality depends on how experts can identify different movements through them. In particular, in order to use them correctly, experts have to perform calculations on them that take the form of a ‘diagram chase’. (4) CDs are hybrid mathematical displays that present both diagrammatic as well as linguistic elements. The analysis of their nature will contribute to the rejection of sharp dichotomies in the classification of mathematical representations. (5) CDs can play essential roles in mathematical proofs, as well as in other phases of the mathematical enterprise.[3] They contribute in a nonreplaceable way to both mathematical discoveries and justifications.

In the present article, I will discuss epistemological issues concerning diagrammatic reasoning. More specifically, I will analyze a case study which can offer new insight on actual mathematical practices involving diagrams. This work is embedded in the tradition of the *philosophy of mathematical practice*, a recent trend in philosophy of mathematics that aims at broadening the scope of philosophy of mathematics to issues that have to do with mathematics as actually practiced by human agents, and not only focused on ontological and epistemological issues regarding abstract objects.[4] I will present CDs, which are a specific type of diagrams that play essential roles in advanced areas of contemporary mathematics such as homological algebra and category theory. In the present article, I will focus on the use of CDs in homological algebra since, as I will explain, in this field they support a specific type of mathematical reasoning, the ‘diagram chase’.

Much work on diagrammatic reasoning in mathematics has been focused on logical diagrams,[5] geometric diagrams, which ‘depict’ spatial relations,[6] and diagrams in analysis.[7] Nevertheless, mathematicians use other kinds of diagrams as well, which function diagrammatically but at the same time present some linguistic features. Feferman (2012, p. 372) acknowledges the presence of non-geometric diagrams in mathematical proofs:

---

[1] This is a new contribution, not fully addressed in other works concerning CDs; see Giaquinto (2007), Feferman (2012), Halimi (2012) and Weber (2013).

[2] ‘Effective’ is here not meant in technical sense: An effective notation is a notation that allows the practitioners to perform the relevant tasks accurately and quickly.

[3] This does not imply that given a proof involving CDs we could not find a different proof of the same statement without diagrams. Nevertheless, many actual proofs contain CDs and eliminating them would mean to alter the proof in a nontrivial way: We would create another proof, probably deprived of the virtues (such as understandability) of the original one. See Giaquinto (2008, pp. 24–26) for a discussion on criteria of identity for proofs.

[4] Among the first and most important publications in this traditions are the collections edited by Mancosu, Jørgensen, & Pedersen (2005) and Mancosu (2008).

[5] For example, Venn Diagrams or their generalization by Peirce or Shin; see Shin, Lemon, & Mumma (Fall 2013).

[6] For example, Euclidean diagrams, as in Manders (2008).

[7] For example, visualizations in calculus, as in Giaquinto (1994).

 https://doi.org/10.1017/S1755020316000277

"the practice of reliance on diagrams in various ways is still integral to the presentation of mathematical proofs of all sorts, even outside of geometry and analysis." With the present case study, I aim at shedding light on the mathematical practices involving CDs in homological algebra.[8]

In §2, I will present some background material on notations in mathematics. I will argue that the introduction of new notations is a relevant part of mathematics: Effective notations matter for mathematics. In fact, as I will show, they allow experts to externalize the relevant reasoning and to perform specific operations. To do so, I will introduce specific terminology to characterize mathematical notations. I will consider three criteria through which it will be possible to evaluate different notations: *expressiveness*, *calculability*, and *transparency*. Briefly, the expressiveness of a notation reveals which kind of information can be conveyed through its use; calculability concerns the possibility of using the notation to perform specific operations with epistemic aims and transparency evaluates the possibility of exploiting our pre-existing cognitive abilities and acquired expertise in interpreting and using correctly the notation. The last label stems from the idea that a notation is 'transparent' if it can be understood and used directly (in a given context), without much explanation or instructions of use. In order to clarify how this methodology can be applied to specific cases, I will discuss two main examples of mathematical notations: numeric notations and notations in knot theory.

In §3, I will present a case study in homological algebra, a mathematical subfield that developed from algebraic topology. I will specify the mathematical details, but only to the extent that they will help in understanding the case study and its philosophical significance. Two main results will be presented: the *Five Lemma* and the *Fundamental Theorem of Homological Algebra*. As we will see, the proofs of both these results rely heavily on the so-called 'diagram chase' technique.

In §4, I will discuss the case study by listing and examining the five main properties of CDs and their uses mentioned above.

In §5, I will sum up some conclusions and hint at new possible research directions.

§2. Mathematical notations. I will introduce in this section some preliminary remarks on mathematical notations. The study of notations is philosophically significant from different perspectives: both from the traditional standpoint of philosophy of mathematics and from the more recent trend focusing on the practice of mathematics. From a more traditional standpoint, notations have been important in relation to foundational issues. It is enough to think about the centrality of the discussion on mathematical notations at the dawn of modern philosophy of mathematics, with Frege's *Begriffsschrift*.[9] Nevertheless, traditionally notations have been studied focusing exclusively on their syntactic and semantic properties, with a constant eye toward formalization.

From the practice based approach, the importance of notations, as pointed out by Grosholz (2007, p. 23), becomes significant not only with respect to its syntactic and semantic dimension, but also to its *pragmatic dimension*. The present contribution is situated in the tradition of the philosophy of mathematical practice and thus notations will be considered as actually used by practitioners. In one seminal contributions in this tradition, Grosholz (2007) considered different case studies concerning the use of heterogeneous representations in the history of mathematics and the sciences. She states her aim as

[8] Another attempt in the direction of considering diagrams outside geometry is for example the analysis of the role of diagrams in free probability theory by Carter (2010).

[9] See Macbeth (2012a) to a detailed analysis of diagrammatic reasoning in Frege's *Begriffsschrift.*

following, by contrasting it with the Carnapian idea of reducing mathematics to logic through syntax:

> My purpose in this book, by contrast, is to move towards an epistemology that works properly for mathematics by taking into account the pragmatic as well as the syntactic and semantic features of representation in mathematics. Focusing on the pragmatic dimension of mathematical language allows us to see the philosophical interest of useful ambiguity in mathematics, as well as the limits of formalization. (Grosholz, 2007, p. 23)

In line with Grosholz, my present goal is to consider how specific notations in mathematics can be evaluated considering their pragmatic dimension, that is, the way in which they are actually employed in the various phases of the mathematical enterprise. More recently, other authors have considered issues concerning notations. For example, Colyvan (2012, chap. 8) points out that "there is a great deal more philosophical work to be done on understanding and appreciating the important role of notations in the various branches of mathematics."[10] Accordingly, one of my present aims is to shed light on the case of the notation formed by CDs.

I claim that effective notations allow users to perform calculations on them, i.e., they are highly "mathematically tractable."[11] This implies that notations not only record information, but put mathematical reasoning in a material form. Specifically, I will argue that the effectiveness of mathematical notations is rooted on the fact that they externalize the relevant reasoning and enhance cognition.

In order to analyze different mathematical notations, I will first report some concepts that have been used to characterize them. Afterwards, I will identify three main criteria to evaluate the effectiveness of a notation. I will start with two remarks on the adopted terminology. First, I will use the term 'notation' in a broad manner, in order to refer to representational systems of various kinds. Therefore, I will include as notations not only linguistic displays such as numerals[12] and algebraic formulas, but also systems of diagrams, such as knot diagrams. In fact, even if the nature of such displays is diagrammatic and not linguistic, they are used in a formal practice in a codified manner. Indeed, it is the possibility of codification that accounts for the existence of formal diagrammatic systems such as the one proposed by Avigad, Dean, & Mumma (2009) for Euclidean geometry.

Also in the case of CDs, diagrams are used in a formal practice. Even if, as we will see, mathematicians use them as single displays with which to reason, they will still be included as mathematical notations. Loosely, I will include as mathematical notations all that mathematicians write down as a cognitive aid in *stable* and *shared* practices. I aim to exclude squiggles that lone mathematicians draw with various aims, such as to promote concentration, in order to include only those signs that carry mathematical content and that can be shared. I will focus on specific notations that are the constitutive elements of formal practices of proving and, more generally, mathematical reasoning.

---

[10] Other recent works on mathematical notations are Brown (2008, chap. 6) and Macbeth (2012c).

[11] Macbeth (2012c) introduces this expression in order to evaluate mathematical notations, in particular Frege's *Begriffsschrift* and Euclid's diagrams.

[12] Although cognitively we appear to have different systems of representations for numbers (in particular for small natural numbers) that are not purely linguistic but visual-symbolic as well, I am interested here in the written notation constituted by Arabic numerals. For the scope of this article, this can be considered as a linguistic notation. Thanks to Marcus Giaquinto for making me clarify this point.

Second, I will distinguish between linguistic and diagrammatic representations. As it will become clear in the following, this distinction is not sharp and it forms more of a spectrum than a dichotomy. As a first approximation, *linguistic* representations are formed by sequences of arbitrary symbols, whereas *diagrammatic* representations exploit their spatial configurations in an essential way, that is, their two-dimensional disposition contributes to defining their semantics.[13]

I will now introduce some terminology developed by Macbeth (2012c) while analyzing different notations such as the one formed by Euclidean diagrams and Frege's *Begriffsschrift*. She introduces the distinction between "trans-configurational" and "intra-configurational" notations. On the one hand, a notational system is trans-configurational if it makes use of "rewriting." Examples of trans-configurational systems are algebraic notations, since when manipulating an algebraic expression, practitioners rewrite equivalent expressions in different ways. On the other hand, a notational system is intra-configurational if it does not require rewriting. As we will see, CDs are intra-configurational. Another case of intra-configurational displays are Euclidean diagrams. In Euclidean geometry, a diagram is a display which presents in an unique figure all the necessary diagrammatic elements for a particular argument: We do not draw different diagrams at each construction step, but analyse the same diagram in different ways.[14]

Another important distinction that Macbeth (2012b) introduces is between "describing" a particular mathematical reasoning and "displaying" it. Macbeth observes that often in mathematics the reasoning is *described* in natural language. This is done through expressions like 'let us consider the composition of...', 'from hypothesis, it follows...' or the description of a mathematical operation. Nevertheless, with the aid of specific mathematical notations, it is possible to *display*, rather than to *describe*, the reasoning:[15]

> The paper and pencil calculation is radically different. It is manifestly not a *description* of a chain of reasoning one might undertake; rather it *shows* a calculation. (Macbeth, 2012b, p. 33)

The crucial observation that a notation can enable us to reason directly within the system of signs is also supported by Krämer (2003), who takes a step further. Krämer analyzed different notations in mathematics and investigated a kind of writing labelled 'operative':

> The advantage of conceiving writing as nonphonetic reveals a whole new realm of written phenomena, which will be called *operative writing* in contradistinction to phonetic writing. Calculus is the incarnation of operative writing. (Krämer, 2003, p. 522)

13 Of course, linguistic representations are linear and thus in a sense also exploit their spatial disposition in an essential way, but in a much more rigid and limited fashion.

14 Alternatively, one can conceive the construction of an Euclidean diagram as a step-by-step process that consists in adding new elements to the same diagram.

15 A mixture of displaying and describing is often present in mathematical notations. In fact, we can adopt specific notations that display the reasoning, and then 'decorate' them with a description in order to make them more transparent. For example, we can use the symbol '$\iff$' to indicate an 'if and only if' relation, as in the expression: $a = b \iff b = a$. If this relation is grounded, as in this case, on the symmetric property of the identity relation, we can indicate this description as follows: $a = b \overset{symm.}{\iff} b = a$. Thus, we mix notations that describes the reasoning, with notations that displays it. The symbol $\iff$ displays the reasoning since it is directly mathematically tractable. For example, we can, without any reference to the meaning of the symbols, invert what is on the left with what is on the right of it.

This notion regards the paper and pencil dimension of signs. Krämer opposes this type of writing to phonetic writing in so far as it is not subject to the linearity of time (not being a simple transcription of speech), but it exploits the two-dimensionality of the paper. Operative writing allows for a process of "de-semantification," that is, we stop paying attention to the meaning of the sign and perform 'blind' operations on them. Through de-semantification we abstract from the semantic dimension and thus forget all meaning.[16] Specifically, to draw inferences with operative writing means to stop following intuitions, because the unique task becomes to apply mechanically the rules defined within the formalism.[17] Therefore, operative writing allows to lighten the cognitive load of the task and to reason directly in the system of symbols. This kind of writing can be observed in various areas of mathematics and can be used in order to better understand certain diagrammatic notations as well. Operative writing is possible only through a notation which, using Macbeth's distinction, displays the reasoning. Nevertheless, these two notions do not coincide. In fact, in the case of operative writing, we not only reason within the system of symbols, but also we do so in an 'automatic' way (though a process of de-semantification).

These distinctions are helpful in clarifying certain characteristics of different mathematical notations. I aim now at identifying three different criteria to evaluate different notations: (i) *expressiveness*, (ii) *calculability*, and (iii) *transparency*. Expressiveness measures which kind of information can be expressed through the specific notation. Calculability determines which calculations can be carried by the notation. In Macbeth's terminology, it expresses which kind of mathematical reasoning can be displayed through the notation. Transparency quantifies how 'intuitive' a notation is. That is, to what extent it can be interpreted and used directly, by exploiting our cognitive abilities and our training. Therefore, the transparecny criterion cannot be fixed once and for all, but it is indexed to the practitioner's background. In fact, it depends on pre-existing cognitive abilities and on the degree of expertise and familiarity with other notations possessed by the practitioner.

In the following discussion, I will refer to these criteria by using them as yardsticks against which to measure different mathematical notations. Nevertheless, I will not propose an algorithm to evaluate notations abstractly. In fact, these three features are present in all notations to different degrees and it is not possible to compare them out of context. Indeed, the effectiveness of a notation depends on its uses.[18] For example, a notation designed with pedagogical aims will be evaluated differently than a notation used for research purposes. Thus, in evaluating a notation in mathematics it will be necessary to consider its pragmatic dimension as well, and not solely its syntactic and semantic nature.

Larkin & Simon (1987) identified two criteria similar to the first two mentioned above. In order to explain the effectiveness of diagrammatic representations in comparison with sentential ones, they evaluated notations in terms of "informational efficiency" and "computational efficiency." The third criterion can be linked to the more general one of "naturalness," as defined in Giardino & Greenberg (2015). According to the authors, a system

[16] Dutilth-Novaes (2012, p. 12) distinguishes three ways in which we generally characterize formal languages as abstracting from content, one of which is the formal as de-semantification (the other two being "formal as topic-neutrality" and "formal as abstraction from intentional content"). She writes: "In this sense, to be purely formal amounts to viewing the symbols as blueprints (inscriptions) with no meaning at all—as pure mathematical objects and thus no longer as 'signs' properly speaking." (Dutilh-Novaes, 2012, p. 13).

[17] This notion is reminiscent of Leibniz's *Ars Combinatoria*, see Ishiguro (1990, chap. III).

[18] The uses of a notation can be the ones intended at its creation but also unexpected uses which might develop afterwards.

 https://doi.org/10.1017/S1755020316000277

of representation "is more or less *natural* to the degree to which human nature—including relatively universal aspects of cognition, physiology, social behaviour, and environmental interaction—rather than enculturation, makes that system easy to internalize and use" (Giardino & Greenberg, 2015, p. 8). I choose a different terminology in order to include both nature and nurture in the evaluation of how transparent is a notation, and not to enter in the discussion of this controversial distinction. From my perspective, the transparency of a notation does not depend only on innate cognitive and psychological characteristics of human nature, but also on cultural factors, such as the habit to see and manipulate a specific notation in a certain way.

Often, expressiveness and transparency are present in complementary degrees. In fact, representations in mathematics present different degrees of abstraction and there is a trade-off between expressiveness and transparency. In a previous work in collaboration with Giardino (De Toffoli & Giardino, 2015), we argued that this phenomenon is very clear in the case of logical diagrams: from Euler diagrams to their generalizations by Venn and consequently by Pierce and Shin. Euler diagrams are very intuitive and easily interpreted, but their expressive power is limited. Venn created new conventions to increase this expressive power, but at the cost of the intuitive interpretation.[19]

In the following, I will present two examples for comparing different mathematical notations: the fist concerns numerical notations (and thus linguistic displays) while the second concerns knot diagrams (and thus diagrammatic displays). Numerical notations are a clear example of how different systems of symbols support different calculations. Arabic numerals have replaced Roman numerals because we can *calculate* with them in a much more efficient way. This is possible because Arabic numerals externalize the mathematical properties of the natural numbers of being recursive, see Brown (2008, p. 85).[20] In order to perform a numerical operation, like a multiplication, we can rely on a specific procedure which exploits the particular properties of Arabic numerals, like in Figure 1.

Once the appropriate notation is established (and with it, its possible manipulations), it allows us to get results 'automatically'. We do not have to think about the specific procedures, as it is an instance of operative writing: The notation carries for us the cognitive load. In the case of Roman numerals (without an algorithm), we would have had to 'think' about the meaning of a multiplication in order to get the desired result. Moreover, in Figure 1, the mathematical reasoning behind the multiplication is externalized and has acquired material form, it is not merely a description of it.

$$\begin{array}{r} 4\,2\,4 \\ \times \quad 2\,4 \\ \hline 1\,6\,9\,6 \\ 8\,4\,8 \\ \hline 1\,0\,1\,7\,6 \end{array}$$

Fig. 1. Multiplication with Arabic numerals.

[19] A description of these different diagrams in logic can be found in Shin et al., (Fall 2013) and it has been discussed in Giardino & Greenberg (2015).

[20] Even if many scholars have accepted this claim, Schlimm & Neth (2008) have cast doubts on its foundations. According to their view, also Roman numerals allow for efficient calculations. Anyhow, this would not weaken my point. Arabic numerals are still superior to Roman numerals in various computational respects. Moreover, both Romans and Arabic numerals are 'better' than a linguistic description in natural language.

Let us compare these two notations adopting the criteria we discussed: (i) the two notations are informationally equivalent, that is, we can express the same information.[21] (ii) We can calculate with both systems much better than with strokes alone or in natural language. Nevertheless, in the case of Arabic numerals, as we have seen, calculations become automatic and are externalized by the notation itself, whereas in the case of Roman numerals, performing arithmetic operations is more cognitively taxing.[22] (iii) Both systems exhibit conventional elements which require a linguistically defined context of interpretation. Nevertheless, in the case of Roman numerals, we have more intuitive elements, such as the strokes for the numbers from one to four. Still, in our culture, we are so familiar with Arabic numerals that it is very easy for us to interpret and use them correctly.

In the actual practice, the differences in the ways we calculate with various notations are not so sharp. In fact, we observe different uses of the same notation. For example, if we have to perform a multiplication, we can do it automatically within the symbols, or first simplify it, by thinking about it in a mathematically pertinent way, and then perform certain tasks automatically. For example, if we want to multiply 424 times 5, we can observe that this is equivalent to performing two more simple operations: first, multiplying 424 times 10 and second, diving it by 2; then we can perform the simpler operations 'automatically' in the notation (for example, the first one would correspond to add a zero). Therefore, there are cases of 'localized automaticity': It is possible to go back and forth between operative writing (through a process of de-semantification) and regular semantic reasoning (with the aid of adding meaning to the symbols).[23]

Macbeth (2012b, p. 60) compares these two numerical notations and claims that only in the case of the Arabic numerals, the collections of signs "do not merely record results; they actually embody the relevant bits of mathematical reasoning." In her view, a mathematical notation is effective only if it embodies the relevant reasoning, by displaying it. Whereas I agree with Macbeth in identifying an essential feature of an effective notation in the embodiment (or externalization) of the relevant reasoning, I disagree that in the case of numeric notations (and more in general as well) we observe such sharp differences. As can be observed from the comparison above, the question is not only whether or not a notation externalizes the mathematical reasoning, but also what exactly does it externalize and in which form (how highly mathematical tractable is the externalization). In fact, as noted above, also Roman numerals may externalize certain properties, just not as well as Arabic numerals. A similar phenomenon, in different degrees, is present with all kinds of notations. A notation will always enable us to lighten our cognitive load by allowing us to use particular representations to reason; therefore, a notation will always externalize the reasoning to a certain extent. The question becomes a matter of to what degree and in which form. With the example of numerical notations in mind, we can identify two extremes. At one extreme, to write an expression in natural language, already aids us (compared to having to memorize it). Indeed, notwithstanding the fact that performing numerical

[21] Nevertheless, the standard Roman system allows to represent numbers only until 4,999, and a generalized version of it until 4,999,999 (Schlimm & Neth, 2008, p. 2097), but of course they can be indefinitely generalized.

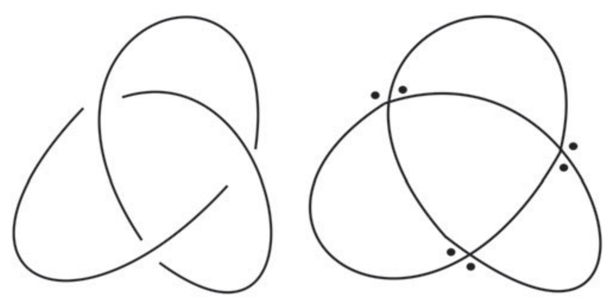
[22] This could be avoided by the adoption of specific artefacts, such as the abacus. Moreover, it has been pointed at computational possibilities allowed by Roman numerals that are generally ignored, see Schlimm & Neth (2008).

[23] This back and forth between an automatic and a semantic type of reasoning is a similar to a phenomenon analyzed by Dutilh-Novaes (2012) and called "de-semantification," as in Krämer (2003), and "re-semantification."

operations without the aid of numerical notations is very hard, already the possibility to write down number words in natural language allows us to externally fix the data and thus let us unload the cognitive task of our memory. At the other extreme, to write a calculation in a system of symbols that are highly mathematically tractable allows us to perform operations that would be inconceivable in natural language.

Let us now shift the discussion to diagrammatic displays. Knot diagrams constitute an effective mathematical notation which functions diagrammatically, but can also be integrated through linguistic displays.[24] Knot diagrams are two-dimensional displays that are easily interpreted as three-dimensional knots. Notwithstanding their pictorial features, knot diagrams do not only display knots, but they are also mathematically tractable, that is, we can perform calculations with them. I will now compare two diagrammatic notations used to denote knots: standard knot diagrams and 'pointed knot diagrams'. This last one makes use of a nonstandard convention at crossings.[25] In Figure 2, the trefoil knot is represented in the two notations.

As I will discuss, these two notations present different degrees of transparency. In fact, Figure 2(a) is immediately interpreted as a three-dimensional curve, whereas in order to interpret correctly Figure 2(b) we need to know that at the crossing the line that goes between the two points is interpreted as going above the other. I will present now simple calculations involving the standard notation for knot diagrams (as I shall argue, this notation is more transparent and therefore the calculations that I will present are understood more easily adopting it). In Figure 3(a), a complicated diagram gets simplified into diagrams with less and less crossings, to end with a diagram with no crossings at all, Figure 3(g). This result is achieved through moves that are easily interpreted using "manipulative imagination," that is, a transposition of the imagination we would actually use for manipulating concrete objects, see De Toffoli & Giardino (2014).[26] Part of the reason why knot diagrams are a transparent notation is exactly that they immediately trigger this type of imagination, which gets enhanced by expertise.



(a) (b)

Fig. 2. Diagrams of the trefoil knot.

24 For a detailed discussion of knot diagrams, see De Toffoli & Giardino (2014).

25 This convention was used by the first knot theorists, for example in Alexander (1928).

26 This is a concept that Giardino and I introduced in a previous work in order to describe how experts use knot diagrams: "the dynamic nature of knot diagrams involves a form of manipulative imagination that gets enhanced through training by transposing our manipulative capacities from concrete objects to this notation." (De Toffoli & Giardino, 2014, p. 839).

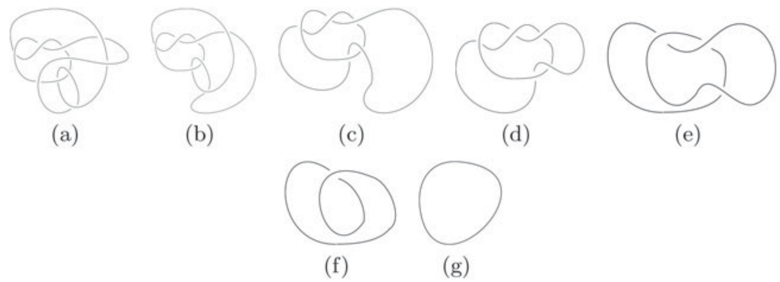


Fig. 3. Diagrams of the unknot.

Looking at the diagram in Figure 3(a) and the one in Figure 3(b), it can be seen that they represent the same knot: It is enough to imagine 'pulling' the middle strand down in order to pass from one to the other. To conclude that all these diagrams also represent the same knot, we just need to apply further similar moves.

Brown (2008, chap. 6) claims that knot diagrams form an effective mathematical notation since we can use them to calculate. Knot diagrams allow for calculations by manipulations (as the ones in the sequence of Figure 3) and, as we will see, through the definition of invariants such as knot polynomials. By this brief description already some properties of knot diagrams emerge. First, they are trans-configurational. In fact, we make use of rewriting when performing an operation with them. Second, they can be used to calculate, that is, they do not only function as descriptions but also display the relevant reasoning. This phenomenon becomes even more clear in the case of knot polynomials. Note that Macbeth does not appreciate the full range of possibilities for calculations offered by knot diagrams, by focusing exclusively on their pictorial features. She claims:

> Diagrams in knot theory [...] directly picture knots that can then be manipulated, in Reidemeister moves,[27] essentially as one would manipulate an actual knot. In all these cases, as in the case of Roman numeration, one directly pictures something and then can manipulate the picture as one might manipulate that which it pictures. A Euclidean diagram, we will see, is different insofar as it (like a numeral of Arabic numeration) does not merely picture something but instead formulates the content of something—in this case, the contents of concepts of various plane figures—in a mathematically tractable way, in a way that enables reasoning in the system of signs. (Macbeth, 2012c, p. 63)

When we consider the rules for the calculation of knot polynomials, as in Figure 4[28] it becomes clear that knot diagrams are much richer than what Macbeth claims. In the calculation of Figure 4, the brackets represent the bracket polynomial operation. In this example, are showed the three rules adopted to calculate the bracket polynomial: (1) The polynomial associated to the trivial knot is just one; (2) If we add a trivial component to a link (i.e., a union of knots), then we have to multiply a factor $(-A^2 - A^{-2})$ to our polynomial; and (3) for each crossing, it is shown how to decompose it.

[27] The Reidemeister moves are local diagrammatic moves that allow one to connect all diagrams representing the same knot.

[28] See Adams (1994, chap. 6) for a discussion and examples of actual computations of this polynomial.

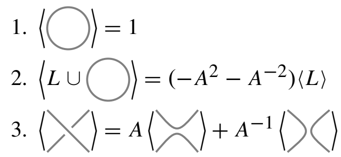


Fig. 4. Rules for the calculation of the bracket polynomial.

The use of knot diagrams as part of the calculations of knot polynomials proves that we do not only manipulate knot diagrams (or parts of them) as we would manipulate actual knots. Through them we can perform different operations which are defined by specific mathematical aims.[29]

I will now evaluate the two notations I presented for knot diagrams in terms of the three criteria identified above. (i) Through them we can represent any knot; they are completely equivalent regarding expressiveness. (ii) We can calculate with them, as in Figure 3, or with knot polynomials. We can perform the same calculations in an equivalent way with the two notations. So they are also equivalent with respect to calculability. (iii) The two notations are clearly not equivalent from the point of view of transparency. In fact, standard knot diagrams are easily interpreted as representing a three-dimensional knot. This is thanks to the conventions used at crossings: The interrupted lines are easily interpreted as 'going under'.[30] This is not the case for pointed knot diagrams: These need a much more explicit set of instructions for interpretation and use. For example, the calculation in Figure 3 would have been much harder to follow in the pointed notation.

In this brief excursus, I showed how mathematical notations are heterogeneous. For example, they can be linguistic, as in the case of numerical notations, or present diagrammatical features, like in the case of knot diagrams. Nevertheless, in order to be effective they must have something in common. In fact, they allow mathematicians to reason *within* the notations; these do not only record information, but also externalize the relevant reasoning, by displaying the mathematical content in a form that is at the same time understandable and tractable. It is therefore possible to calculate through an effective notation. Thus, there are not specific characteristics to follow for a particular notation to be effective, since the aim of externalizing mathematical content can be obtained in different ways. The three criteria identified above are useful to understand the functioning and the effectiveness of particular notations, but we have to remember that to evaluate a notation we must consider its intended context of use. A general observation is that, if in two notations two criteria are equal and one can vary, then the notation which is superior according to this third criterion would be preferable. For example, as we saw, in the case of notations for knot diagrams, the standard and the pointed notations differ only by their transparency. Thus, the notation that is more transparent, which is the standard one, is preferable.

In the following sections, I will analyze how CDs form a good mathematical notation in homological algebra, which present features proper to diagrammatic as well as linguistic representations. I will maintain that the specific practices involving CDs, as well as CDs themselves, are epistemologically relevant, since they are integral parts both of the reasoning and of the justification provided. As we shall see, CDs, like other effective notations, do not simply record information, but externalize the relevant reasoning.

---

[29] A similar argument can be deployed for the case of Roman numerals as well. Through devices such as the abacus we can manipulate the symbols in order to obtain valid mathematical results, as explained by Schlimm & Neth (2008, p. 2097).

[30] In fact, standard knot diagrams exploit certain laws identified by Gestalt psychology, in this case the law of "good continuation" of visual perception, see Kanizsa (1980).

§3. Case study. The main goal of algebraic topology is to apply techniques and results of abstract algebra to topological inquiries. Algebraic invariants are among its most important tools. To topological spaces we associate algebraic structures that do not change if we replace the given space with one homeomorphic to it. The converse does not hold: In general the same algebraic structure can be associated to different topological spaces.[31] Nevertheless, invariants are very useful because they allow one to discriminate, and in some cases to classify, topological spaces while working with more mathematically accessible objects such as groups or other algebraic structures.

A very important algebraic invariant is *homology*, which will be defined below. Even if homology was first introduced in order to study topological spaces, it became interesting on its own. In fact, now there exists an entire branch of algebraic topology, *homological algebra*, which deals with the homology of algebraic structures, oblivious of the original topological motivations.[32]

3.1. Commutative diagrams. Let us start with some basic definitions.

DEFINITION 3.1. *A* chain complex $(C_\bullet, \partial_\bullet)$ *is a sequence of abelian* (*i.e., commutative*) *groups connected by homomorphisms:*

$$\cdots \overset{\partial_{n+1}}{\to} C_n \overset{\partial_n}{\to} C_{n-1} \overset{\partial_{n-1}}{\to} \cdots \overset{\partial_2}{\to} C_1 \overset{\partial_1}{\to} C_0 \overset{\partial_0}{\to} C_{-1} \overset{\partial_{-1}}{\to} \cdots .$$

The homomorphisms $\partial_\bullet$ *are called* boundary maps.[33] *The composition of two consecutive homomorphisms is always the zero map:* $\partial_{n-1}\partial_n = 0$[34] *for all* $n \in \mathbb{Z}$.

In order to define homology groups, we need to analyse the images and the kernels of the boundary maps. In general, the *image* of a homomorphism $f : X \to Y$ is the subset of Y defined as $\mathrm{Im}(f) := \{f(x) \in Y | x \in X\}$. The *kernel* of f is the subset of X made of all the elements which get sent to the zero element: $\ker(f) := \{x \in X | f(x) = 0\}$.

Note that if we consider the boundary operators, we have that both $\mathrm{Im}(\partial_{n+1})$ and $\ker(\partial_n)$ are subsets of C_n. Moreover, since for all n, $\partial_n\partial_{n+1} = 0$, we have that $\mathrm{Im}(\partial_{n+1}) \subseteq \ker(\partial_n)$. We can therefore define the homology groups as follows:

DEFINITION 3.2. *The n-th* homology group *of a chain complex* $(C_\bullet, \partial_\bullet)$ *is the quotient:*

$$H_n := \ker(\partial_n)/Im(\partial_{n+1}).$$

We then characterize chain complexes according to their homology:

DEFINITION 3.3. *A chain complex is* exact *at* C_n *if* $H_n = 0$.

Therefore, a sequence is exact at C_n if $\ker(\partial_n) = \mathrm{Im}(\partial_{n+1})$. We say that a sequence is exact if it is exact at each node. Homology, in purely algebraic terms, basically measures how far a sequence is from being exact.

31 The converse only holds for the so called *universal invariant*.

32 For a rigorous definition of homology see Bredon (1993). As reference manual for algebraic material, including groups, homomorphisms, short exact sequences, and CDs, see Lang (2002).

33 The reason why the homomorphisms are called *boundary maps* is to be traced to the geometric origin of these algebraic objects. More specifically, a topological space can be encoded into a chain complex thorough a *cellular decomposition*. Then, the homomorphisms are actually the *boundary* maps which define how the cells are attached to each other through their boundaries.

34 It is common in the practice to indicate with '0' different mathematical objects: the trivial group, the trivial element, and the trivial homomorphism, sending all elements to the trivial element and the empty set. Thanks to Marcus Giaquinto for pointing out this ambiguity to me.

 https://doi.org/10.1017/S1755020316000277

We introduce a particular type of sequence, as it will be relevant in the following:

DEFINITION 3.4. *A* Short Exact Sequence (*SES*) *is a sequence that is exact everywhere and has the following form:*

$$0 \to A \xrightarrow{f} B \xrightarrow{g} C \to 0.$$

Given a SES, the following properties follow from the definition:

- f is injective. In fact, by exactness at A and observing that any homomorphism from the trivial group must be itself trivial, we have that $\ker(f) = \mathrm{Im}(0) = 0$. (By definition, a map is injective when the kernel is zero.)
- g is surjective. In fact, $\ker(0) = \mathrm{Im}(g)$, but the kernel of the trivial map is the whole domain, so $\ker(0) = C$.
- $\ker(g) = \mathrm{Im}(f)$.

Let us now consider two-dimensional 'arrow diagrams', and not only sequential ones:

DEFINITION 3.5. *A* commutative diagram *is composed by nodes and arrows connecting them. The nodes are a certain type of mathematical objects and the arrows are morphisms connecting them. A diagram is commutative if any two paths connecting a node to another one are equivalent.*

For example, the diagram in Figure 5 is commutative if and only if following the two paths from A_1 to B_2 we get the same result, that is, $g \circ \alpha = \beta \circ f$. In terms of elements, if for every $a \in A_1$, then $g(\alpha(a)) = \beta(f(a))$.

In particular, we can consider the case where the objects are abelian groups and the morphisms are homomorphisms. The study of commutative diagrams in general is part of category theory.[35]

CDs are extensively used in algebra. In fact, they serve to display in a single representation different algebraic information. They are frequently used to represent basic properties of modules or of groups, see for example Lang (2002, chap. 1). For instance, we can represent properties of quotient spaces in a single representation as in Figure 6.[36] The mathematical details are not relevant here, I just aim to show different examples of CDs.

I will now present a last example of a CD, the one in Figure 7, before getting to the diagram chase technique. This diagram appears in the proof of the Van Kampen Theorem,

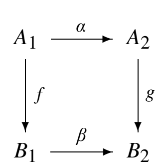


Fig. 5. A CD.

[35] See Simmons (2011) for an introduction to category theory.

[36] The diagram in Figure 6 refers to the following theorem:

THEOREM 3.6. *Let G be a group and N one of its normal subgroups. Let $f : G \to H$ be a group homomorphism such that $N \subset \ker(f)$. Then, there exists a homomorphism $f_* : G/N \to H$ such that the diagram in Figure 6 is commutative. That is, we have $f = f_* \circ \phi$.*

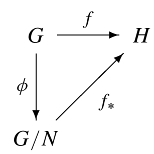


Fig. 6. A triangular CD.

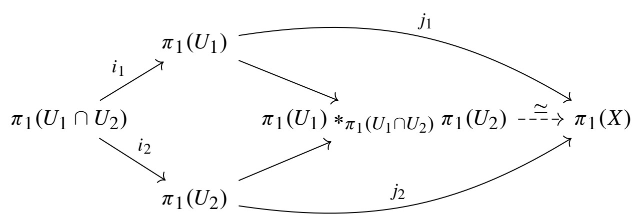


Fig. 7. An example of a CD.

see Bredon (1993, pp. 159–160). Again, my aim here is just to show that CDs can assume different forms and help in proving various mathematical propositions, the technical framework and details are not relevant in this context.

In these last two examples, CDs help to describe an algebraic situation and to infer results about it. Although also in these cases they are a great aid for mathematicians, the diagram chase technique is not involved. Therefore, let us now turn to CDs that support the diagram chase technique.

3.2. Diagram chase. I will present two proofs in homological algebra that make use of the so-called 'diagram chase' technique. I will argue that this technique involves representations which display both linguistic and diagrammatic features. As we will see, these proofs are based on the analysis of CDs. I will claim that these diagrams play different roles and are indispensable in understanding the relevant reasoning behind the proofs and even in phrasing in a meaningful way the statements of the theorems.

One major difference between the two proofs is that the first involves a diagram which is a finite display of nodes and arrows, while the second involves a diagram which can be extended indefinitely: Suspension points are used at its extremities in order to indicate an arbitrary continuation. In this last case, I will interpret the diagram itself as the finite display, but the suspension points play a crucial role in its interpretation, as we shall explore in the discussion. There are deep philosophical issues that emerge in considering this type of 'infinite diagrams'. In particular, there are questions concerning the possibility of reasoning with diagrams by induction. Although this and other connected issues are relevant to the present study I am obliged to postpone their treatment to further work. To get an idea of issues that emerge considering infinite diagrams, from the point of the possibility of formalizing them, see Feferman (2012).

I hope to convey an idea of how proofs by diagram chase go through, without discussing technical details which would be beside the point in this context.

3.2.1. The Five Lemma.

LEMMA 3.7 (Strong version of the Five Lemma). *Let the diagram in Figure 8 be commutative and such that its rows are exact. If f_2 and f_4 are surjective and f_5 is injective, then f_3 is surjective. Symmetrically, if f_2 and f_4 are injective, and f_1 is surjective, then f_3 is injective.*

$$\begin{array}{ccccccccc}
A_1 & \xrightarrow{\alpha_1} & A_2 & \xrightarrow{\alpha_2} & A_3 & \xrightarrow{\alpha_3} & A_4 & \xrightarrow{\alpha_4} & A_5 \\
\downarrow{\scriptstyle f_1} & & \downarrow{\scriptstyle f_2} & & \downarrow{\scriptstyle f_3} & & \downarrow{\scriptstyle f_4} & & \downarrow{\scriptstyle f_5} \\
B_1 & \xrightarrow{\beta_1} & B_2 & \xrightarrow{\beta_2} & B_3 & \xrightarrow{\beta_3} & B_4 & \xrightarrow{\beta_4} & B_5
\end{array}$$

Fig. 8. CD for the 5-Lemma.

Proof of the first part. I will prove the first part of the Lemma. The proof of the symmetric second part deploys the same technique and presents a similar structure, and will be omitted here.

The proof consists in 'chasing' the diagram, that is, in moving an element through different paths of the diagram, as it is shown in Figure 9.

We want to prove that f_3 is surjective, given that f_2 and f_4 are surjective and f_5 is injective. In other words we need to show that the image of f_3 is all B_3: $\mathrm{Im} f_3 = B_3$. So, let us start with an element $b_3 \in B_3$, as in Figure 9(a). We want to show that there is an element $a \in A_3$ such that $f_3(a) = b_3$. Now, we can carry b_3 around the diagram until we find such an a. Let $b_4 := \beta_3(b_3)$, as in Figure 9(b). Since f_4 is surjective, there is an element $a_4 \in A_4$ such that $f_4(a_4) = b_4$, as in Figure 9(c). Now, $\beta_4 \circ \beta_3 = 0$ since the rows form exact sequences.[37] We have: $\beta_4(b_4) = \beta_4(\beta_3(b_3)) = 0$, as in Figure 9(d). Since f_5 is injective, the preimage[38] under f_5 of the zero element, is zero. Moreover, the diagram is commutative: $f_5(\alpha_4(a_4)) = \beta_4(f_4(a_4))$. But $\beta_4(f_4(a_4)) = \beta_4(b_4) = 0$, therefore $f_5(\alpha_4(a_4)) = 0$. Then, $\alpha_4(a_4) = 0$, as in Figure 9(e).

That is, $a_4 \in \ker \alpha_4$. Since the rows are exact, we have $\ker \alpha_4 = \mathrm{Im}\alpha_3$. Then, there exists $a_3 \in A_3$ such that $\alpha_3(a_3) = a_4$, as in Figure 9(f). Now, if $f_3(a_3) = b_3$ we would be done, but we cannot assume it. Let us first suppose that $f_3(a_3) = b_3'$; we have $\beta_3(b_3 - b_3') = \beta_3(b_3) - \beta_3(b_3') = b_4 - b_4 = 0$. But then we can suppose that $\beta_3(b_3) = 0$; if this is not the case we just subtract from b_3 the element b_3' which has a preimage in A_3.[39] Then, $b_3 \in \ker \beta_3$. But $\ker \beta_3 = \mathrm{Im}\beta_2$ by exactness. So there exists b_2 such that $\beta_2(b_2) = b_3$. Then, since by hypothesis f_2 is surjective, we can consider a preimage of b_2 under f_2 and find a_2 such that $f_2(a_2) = b_2$. We have, $\alpha_2(a_2) = a_3$. Therefore, by commutativity, $f_3(\alpha_2(a_2)) = \beta_2(f_2(a_2))$. That is, $f_3(\alpha_2(a_2)) = f_3(a_3) = b_3$, as wanted. □

[37] Actually, for this relation to hold we just need that the rows form chain complexes.

[38] A preimage of an element b under a homomorphim $f : A \to B$ is an element $a \in A$ such that $f(a) = b$. Of course, it does not always exist. A preimage of b exists only if $b \in \mathrm{Im} f$. That is, we follow the arrow in the inverse direction.

[39] These are standard moves in homological algebra, although they might be opaque to the unfamiliar reader. More explicitly, we wanted to prove the existence of an element a such that $f_3(a) = b_3$. If $f_3(a_3) = b_3$, then we are done. Otherwise, suppose that $f_3(a_3) = b_3'$, then $\beta_3(b_3') = \beta_3(f_3(a_3)) = f_4(\alpha_3(a_3))$ by commutativity of the diagram. But $\alpha_3(a_3) = a_4$ and $f_4(a_4) = b_4$. Therefore, $\beta_3(b_3') = b_4$. But we also have that $\beta_3(b_3) = b_4$. Thus, $\beta_3(b_3 - b_3') = \beta_3(b_3) - \beta_3(b_3')$ (since β_3 is a group homomorphism) which is equal to $b_4 - b_4 = 0$. Now we can consider to substitute b_3 with $b := b_3 - b_3'$. This will not alter the generality of the result since b_3' is an element in B_3 that has a preimage. We wanted to prove that b_3 had a preimage, now, if we add or subtract an element with a preimage to it, nothing changes: We modified the element b_3 in a trivial way, with respect to what we wanted to prove. In the practice, thanks to the specific display, these steps are automatic (or semiautomatic).

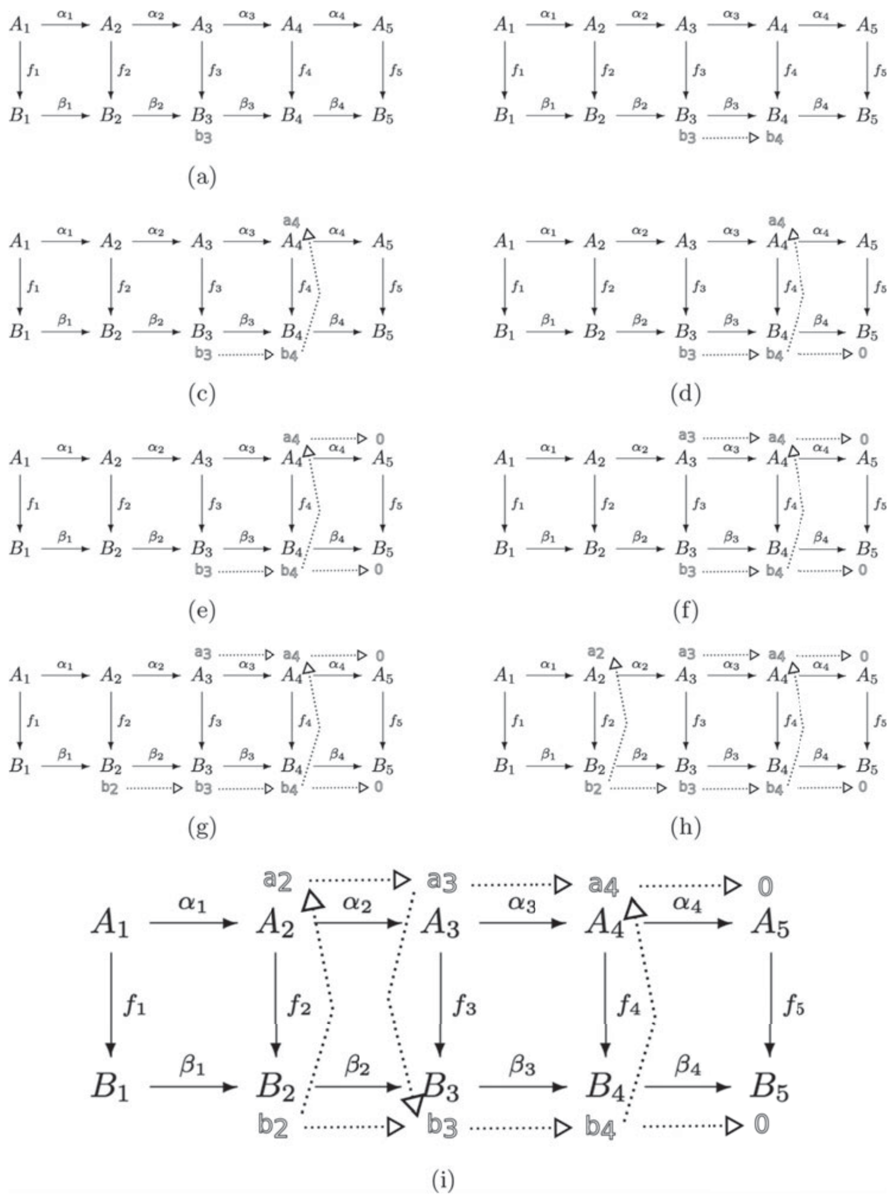


Fig. 9. Display of a "diagram chase."

3.2.2. FTHA. In this section, I will briefly present the Fundamental Theorem of Homological Algebra (FTHA).[40] As I will hint, the proof is based on a diagram chase in a diagram that is potentially infinite.

40 For a complete statement and proof see Bredon (1993, p. 178).

THEOREM 3.8 (FTHA). *A short exact sequence* $0 \to A_\bullet \xrightarrow{i} B_\bullet \xrightarrow{j} C_\bullet \to 0$ *of chain complexes induces a long exact sequence on homology*

$$\cdots \to H_{p+1}(C_\bullet) \xrightarrow{\delta_\bullet} H_p(A_\bullet) \xrightarrow{i_\bullet} H_p(B_\bullet) \xrightarrow{j_\bullet} H_p(C_\bullet) \xrightarrow{\delta_\bullet} H_{p-1}(A_\bullet) \to \cdots.$$

As you can see in Figure 10, the rows are all SESs, while the columns are chain complexes; these are indefinitely long, with indices in the integers. The idea of the proof is to "carry elements around in the diagram" (Bredon, 1993, p. 178) in order to appreciate different properties of the various homomorphisms.

A step of the proof. The proof is based on the analysis and use of the commutative diagram in Figure 10. In order to prove this theorem, once again, we have to chase the diagram. That is, we identify different paths exploiting the commutativity of the diagram, in order to verify some algebraic properties.

In particular, we have to connect the two homology groups $H_p(C_\bullet)$ and $H_{p-1}(A_\bullet)$. This is represented in the diagram in Figure 11(a).

I will give a brief idea of how this is mathematically done. First, we will check the existence of a homeomorphism $H_p(C_\bullet) \to H_{p-1}(A_\bullet)$, see Figure 11(a). In order to understand the following reasoning, it is necessary to follow the steps in the diagram represented in Figure 11(b). We have to check that an element $c \in H_p(C_\bullet)$, i.e., such that $\partial c = 0$, is sent to an element in $H_{p-1}(A_\bullet)$, that is an $a \in A_\bullet$ such that $\partial a = 0$. Thus, we consider an element $c \in C_p$ such that $\partial c = 0$. Since j_* is surjective, we can take $b \in B_p$ such that $j(b) = c$. By commutativity, we have (omitting the indices) $j(\partial(b)) = \partial(j(b)) = \partial(c) = 0$. Then, $\partial b = b'$ is in $\ker j$. But $\ker j = \mathrm{Im} i$, since the rows are exact. Then there is an unique $a \in A_{p-1}$ such that $i(a) = b'$. Then, by commutativity we have $i(\partial(a)) = \partial(i(a)) = \partial b' = \partial\partial b = 0$. Therefore, $\partial(a) = 0$, since i is injective. Then, for each $c \in H_p(C_\bullet)$, there is an unique $a \in H_p(A_\bullet)$ as wanted. □

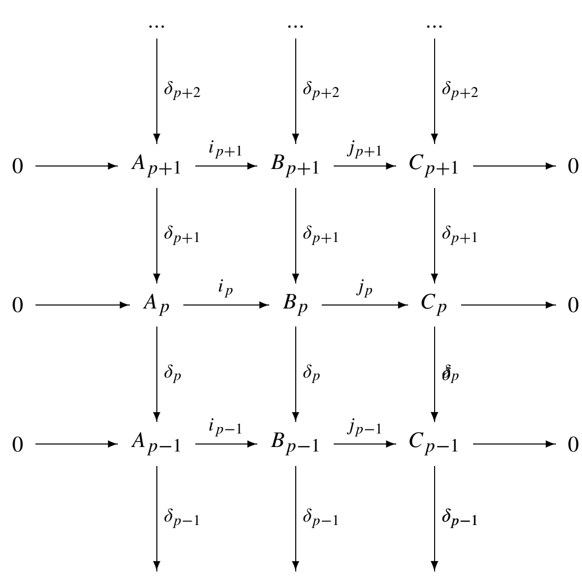


Fig. 10. CD for the FTHA.

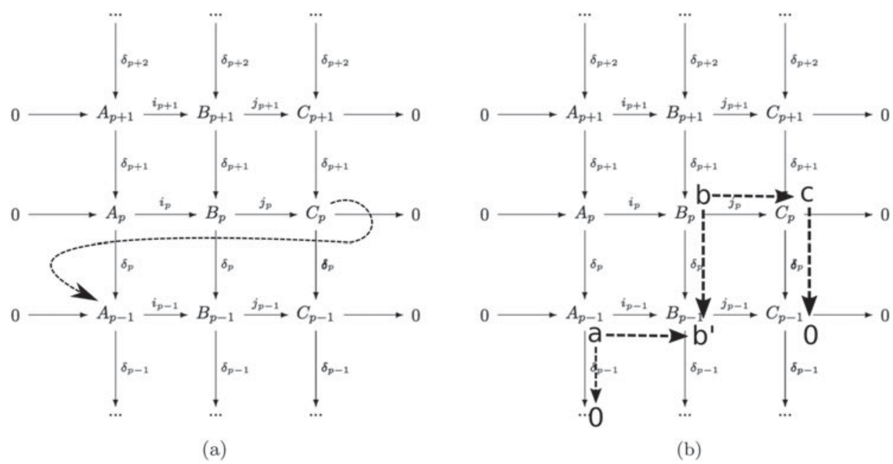


Fig. 11. Step in the proof of the FTHA.

§4. Discussion of the case study. Through the display of the 5-lemma and of the FTHA, it is clear that CDs play an essential role not only in proofs, but also in defining the relevant content-matter in the statement of a theorem. In fact, without diagrams, even stating the results would be difficult. Not only it would be longer, but, as I will discuss, it would be also difficult to express the two results in a meaningful way. In the following, I will focus on the five aspects concerning the use of CDs identified above.

4.1. CDs form an effective notation. CDs are ‘arrow-diagrams’ that depict algebraic relations: Not only do they record all the relevant information, but they make it immediately accessible by representing the whole algebraic structure at once. This makes them effective as a mathematical notation, in the (nontechnical) sense that they function well in the practice in which they are deployed: They offer a convenient way to find a sequence of steps to prove a mathematical result.

In the case of the 5-lemma, the display is complete in the sense that the entire situation relevant for the proof is depicted simultaneously. In the case of the FTHA, the diagram represents a relevant chunk of information—only three rows are displayed—but these are enough to obtain a general result. In fact, in order to prove the result, we choose an arbitrary index p and then analyze the situation in the rows corresponding to the indices from $p-1$ to $p+1$. It is the arbitrariness of p that makes the proof general enough.

Even if these diagrams resemble a standard algebraic notation, spatial elements are exploited to externalize the relevant reasoning. In particular, planar configurations of the diagram correspond to algebraic relations. The arrows are to be interpreted as homomorphisms between groups and the different sequences of arrows, or paths, are compositions of homomorphisms.

All spatial relations are to be interpreted topologically and not metrically. In fact, the exact shape and length of the arrows is irrelevant in any correct interpretation: Only their combinatorial structure matters. For example, in a CD a square is equivalent to a rectangle, but not to a triangle (as the one in Figure 6).[41] In the case of CDs, metric properties do not play a mathematical role.

[41] To further generalize, topologically, any polygon could be continuously deformed into a circle with nodes. Nevertheless, the disposition matters *cognitively*: Using squares and triangles instead

Another aspect of the spatial configuration of CDs is that in the case analyzed, these spatial elements are the disposition of the nodes and arrows in two dimensions to form squares or triangles. Generally, CDs in homological algebra are formed by these two shapes, but they could be in principle also made by other figures, such as pentagons or hexagons.[42] Moreover, theoretically, CDs could have an arbitrary number of arrows touching any given node.[43] Therefore, the fact that CDs in homological algebra present these simple shapes is not a fixed feature intrinsic to their nature, but depends on the specific algebraic structure represented.

The fact that CDs present a simple and standard spatial disposition makes them particularly easy to interpret and use. For example, SESs, and their combinations into squares, are a standard representation of algebraic content. As soon as the groups and homomorphisms are represented in this way, certain properties become evident to the expert. For instance, given a display as the one in Definition 3.4, for an expert it is automatic that f is injective and g is surjective.[44] Therefore, by exploiting the characteristic of standard shapes, mathematicians obtain results from the representations in an automatic way. This is a case of 'localized' operative writing. It is localized because just certain operations are performed automatically, while others require a specific mathematical interpretation and semantic reasoning. This case is similar to the one in which to perform a multiplication an agent simplifies (in a not automatic way) the task into simpler ones that then are to be performed automatically.

Let us now consider how the three aspects we used to analyze mathematical notations (expressiveness, calculability, and transparency) are modulated in the case of CDs: (i) CDs are used to express algebraic relations. As we already observed, they generally come in standard forms (triangles and squares) but they could assume different shapes as well. The possibility to generalize the notation shows that they are highly expressive. Part of their expressiveness derives from the linguistic and abstract components they exhibit. I will develop this point in §4.2 where I will explore the similarities of CDs with geographic maps. In the context of homological algebra, they express any discrete collections of homomorphisms between groups. Nevertheless, the same notation can be used to represent other mathematical categories.

(ii) Mathematicians apply the 'diagram chase' technique through CDs. The types of calculations CDs permit are not standard symbol-manipulations, but involve more conceptual processes. These operations are performed by moving elements around the diagrams. Therefore, the particular form of CDs makes them effective in conveying in a tractable form the relevant mathematics. The nature of these calculations reveals the dynamic nature

of circles, we recognize more easily the corresponding algebraic relations. The notation is thus more transparent, at least for the community of experts.

[42] Here I refer to 'combinatorial shapes', that is, the topological properties of CDs—the issue that instead of squares we could have circles is a distinct one, geometrical in kind.

[43] Of course there is a practical limitation, a CD with nodes that engage a big number of arrows would be hard to interpret and use correctly: There is a cognitive limitation at play.

[44] Note that a first example of CD is just a linear sequence of nodes and arrows, like a SES (for example a row in Figure 10). This is a limit case since in a sequence the commutativity condition is automatically satisfied: there is always a unique path joining two nodes. Nevertheless, the form of the representation in a SES is already relevant. Through it, we visualize the relevant maps all at once and we can infer some properties immediately. When deploying SESs, experts exploit their standard form in order to reason within the system of symbols. It is the specific arrangement in the plane of the paper that allows statement like "the rows are SESs" and "the columns are chain complexes."

of CDs, which I will explore in §4.3 CDs combine two different representational modes. On the one hand, they allow for an automatic 'movement' of elements in the diagram. On the other hand, as I will develop in the following section, they are 'algebraic maps' that present all the relevant algebraic relation in an unique two-dimensional display.

(iii) CDs are highly conventional, so interpretation is needed in order for them to be used correctly. We need to add linguistically conveyed information to the diagrams, in order to interpret and use them correctly. For example, we know that two paths sharing the same starting and ending points are equivalent, that is, we need to be aware of the commutativity condition. Nevertheless, the simple geometric structures they present make them intuitive. The diagrammatic parts (arrow and planar disposition) are easily interpreted as algebraic relations, if the correct context of interpretation is given.

***4.2. CDs function as* maps.** Let us now analyze another feature that contributes to the effectiveness of CDs: their similarities to geographic maps.[45] As I explained, in order to correctly use a CD, not only does an agent need to understand how to interpret it correctly, but she also has to be able to 'move' between its various nodes. These diagrams can be interpreted as *maps*: CDs display the geography of an algebraic structure, where the mathematicians have to get oriented and be able to move abstract elements. In order to use CDs correctly, experts have to interpret the statements they want to prove in terms of the paths represented in the diagrams: A proof by diagram chase consists in moving an element around a map of an 'algebraic world'. A more accurate analogy would be with maps such as the London underground map, which are made of a finite number of interconnected nodes. In this type of maps (presenting a schematic nature), like in CDs, metric information does not really matter,[46] but the relevant information is carried by the combinatorial structure.

As maps, CDs display extra information 'for free', that is, they make available information that is not included in the set of linguistic data used to construct the diagram. Looking at the London underground map, we are immediately aware of how many lines intersect in a certain station or how many stops separate two stations. Analogously, looking at a CD we are immediately aware of all the possible paths from two nodes. Moreover, through extra information such as which maps are surjective, we can know in which cases we can follow an arrow backwards.[47]

As concrete maps allow us to grasp the landscape they represent at first sight, so CDs allow us to understand the whole algebraic situation immediately. Weber (2013, p. 7), considering a CD depicting the algebraic property of being a *free module* (i.e., a specific algebraic structure) writes: "Putting this notation into two dimensions makes a figure; and while the formal definition requires some time to absorb, the picture shows exactly and immediately what it means for F to be free."

As we have seen in the case study, in order to carry out a proof which is based on a diagram chase, we must be guided by the written text step by step and follow in the diagram the various movements described. If we accept the analogy with maps, CDs represent algebraic landscapes, but they do not themselves gives us instruction on how

45 See Winther (Forthcoming) for a philosophical journey into the realm of geographic maps.

46 Of course, it matters *cognitively* to a certain extent. Nevertheless, it does not carry relevant information that can be read off.

47 This extra information can be integrated in the diagram by adopting specific conventions. For example, a two-headed arrow $\twoheadrightarrow$ indicates that the map is surjective and a hooked arrow $\hookrightarrow$ indicates that the map is injective.

to use them. We have to interpret them correctly and 'translate' what we want to prove in terms of paths in the diagrams. In fact, to perform a diagram chase, we need to pay attention at each step of the proof to a different node of the diagram and use the diagram to identify possible moves. This implies that the written text must guide us in the correct use of the diagrams. A similar phenomenon can be observed also in the case of Euclidean diagrams: As Manders (2008) pointed out, reasoning in Euclidean geometry consists in a constant feedback between figures and text. As in the case of geographic maps, not only are text and diagrams in continuous feedback in the case of CDs, but linguistic elements are also integrated in the diagrams. These are for instance indices and letters for CDs; in geographic maps we can find similar entries or even complete words. These linguistic elements are certainly advantageous, but they should not overshadow the importance of the two-dimensional disposition of CDs (and, more generally, of maps).

Another feature that CDs have in common with Euclidean diagrams (and with certain maps) is that they are intra-configurational.[48] Note that, in order to make the diagram chase visible, I rewrote the same diagram many times in Figure 9. Nevertheless this rewriting is normally omitted. The information carried by the extra-diagrammatical elements added to the diagram allowed me to include an encoding of the text into the diagram. Normally experts 'read' the diagram without rewriting. What I drew in Figure 9 is not a sequence of different diagrams, but sequences of *the same* diagram, where the relevant interpretation was made explicit by adding elements that did not strictly belong to the commutative diagram. It would be similar to drawing many diagrams for the same proposition in Euclidean geometry, in order to make explicit the construction of a single diagram, or again, to draw different maps of the London underground to highlight a specific path step-by-step.[49] This aspect has been identified by Halimi (2012, p. 389) who, taking the 5-lemma as an example, states that adding an element does not lead to a new diagram: "We do not have two diagrams here, but just a single one, which is both a starting point and the embodiment of the successive updates involved in its completion." In §4.3, I will analyze in detail in which sense the "updates" can be regarded as an externalization of a mathematical reasoning, which is, as I claim, allowed by the dynamic character of CDs.

In the case of CDs, we can quantify how much the particular disposition in two dimensions is optimal compared to a sequential one. First of all, we need to write sequentially all the maps forming the diagram. This means that if we have two arrows from a node, then in a purely sequential 'translation', we would have to write the node two times to describe both homomorphisms. Then, in order to impose the commutativity, we would have to add an equation for each square and triangle. In fact, a diagram is commutative if the path we follow to connect two nodes is irrelevant, that is, we would obtain the same result by following an alternative one. Thus, we have to check this for each rectangle or other closed cycle.[50]

[48] This is not the case for CDs in category theory, where we might need to draw sequences of diagrams. But we do not need such a rewriting in order to prove a result by diagram chase.

[49] The importance of using a CD step-by-step unveils the importance of the transmission of the construction steps of the diagrams. In the context of a mathematical classroom or seminar, it is much easier to convey in a vivid way this aspect rather than through published material.

[50] Note that the commutativity of a rectangle formed by more than one square follows from the commutativity of the squares constituting it. For example, in the diagram of the 5-lemma, we only need to impose the commutativity of the four small squares and then the commutativity of the whole diagram follows directly.

4.3. CDs are dynamic. The fact that CDs present the relevant material in a unique display, is already a difference between CDs and a list of equations, but the point is not only that a diagram makes us save space. The important aspect is that with diagrams we can access to the whole situation simultaneously and calculate directly through the representations. Through a CD we get an unique display of the homomorphisms involved and their relations. The record of the same information via equations may obscure the overall situation.

Weber (2013, p. 8) analyzes different examples of abstract spatial notations, such as CDs. He claims that "all this pure syntax is not just bookkeeping. It is a method for capturing structure, in compressed and manipulable forms." This assertion is in line with the claim that CDs form an effective notation partly because we can calculate with them.

In the case of CDs, diagrams are not only shorter displays, but they are also essential for us to understand the mathematical situation. We could feed to a computer a sequentially encoded diagram, but we would have to reconstruct it in order to understand the reasoning behind a proof that involves it. The point is that CDs do not just record information, like (in a first approximation) a sequence of equations would, but they display information in a way that makes it accessible and mathematically tractable. By externalizing the reasoning, CDs allow mathematicians to reason within the symbols and to exploit their ability to reason spatially to draw inferences. In this sense it is possible to understand the claim that CDs externalize mathematical reasoning.

Moreover, the fact that CDs support the diagram chase technique reveals their dynamic nature. Chasing the diagram is a specific type of calculation. In order to perform such a calculation, experts interpret CDs as dynamic reasoning tools, rather than static illustrations. In a previous work, in collaboration with Giardino (De Toffoli & Giardino, 2014), we considered the difference between *static* and *dynamic* mathematical representations. We claimed that it is possible for the same figure to be interpreted in different ways, either as a diagrams or as an illustration. For example, if a figure representing a knot is interpreted as a diagram, then it will become a mathematical notation supporting specific calculations. On the contrary, if it is interpreted as an illustration, it will only depict a geometric object.

> By *illustration*, we mean a static representation, which can be useful by conveying information in a single display, but where modifications are not well-defined. By *diagram*, we mean a dynamic representation, on which we can perform moves that can count as inferential procedures. Diagrams are dynamic inferential tools that are modified and reproduced by the experts for various epistemic purposes. (De Toffoli & Giardino, 2014, p. 830)

Adopting this characterization, CDs are dynamic since chasing the diagram is an inferential procedure they support. This calculation requires the ability of focusing on different parts of the diagram and identifying different paths in it. Giaquinto (2007, chap. 12) has identified three types operations in spatial thinking, which he labeled as: "noticing reflection symmetry," "aspect shifting," and "visualizing motion." In the case of CDs, these operations are combined into a new one: *paths identification*, which is the same type of reasoning we perform with geographic maps (or better, with schematic maps such as the one of the London underground). Therefore, CDs present a dynamic nature in which different types of visual operations converge.

The dynamic character of diagrams has been stressed also in the case of diagrams in category theory. In particular, Halimi (2012) has put forward the notion of "evolving diagrams" as diagrams that present a specific dynamic character. According to the author, an evolving diagram would not only represent a mathematical object, but also support a specific mathematical reasoning or construction. CDs and other diagrams present similarities with Euclidean diagrams:

> One of the best examples of the latter feature is the Euclidean figure, which in a proof is often supplemented by the auxiliary lines and circles that the proof calls for. I would like to defend the idea that this is in fact a quite general fact: many mathematical diagrams do not represent mathematical objects once and for all; rather, they can be manipulated and enriched for the purpose of unfolding mathematical properties of the respective objects which they represent. (Halimi, 2012, p. 388)

To return to Macbeth's essential distinction between describing and displaying mathematical content, CDs, via their dynamic character, display rather than describe the reasoning. This in particular implies that CDs are mathematical objects in themselves, and not only a descriptive tool.[51]

4.4. Beyond sharp dichotomies. As we saw, CDs are a hybrid notation in mathematics, containing at the same time diagrammatic as well as linguistic elements. These different elements support different kinds of reasoning, which can be labeled as geometric and algebraic. Nevertheless, as Giaquinto observed, it is misleading to conceive a sharp division between geometric and algebraic reasoning in mathematics:[52]

> We commonly draw a distinction between algebraic thinking and geometric thinking. While that classification may suffice for casual discussion in restricted contexts, I claim that the algebraic–geometric contrast, so far from being a dichotomy, represents something more like a spectrum. To the extent that there is a fundamental dichotomy in mathematical thinking, it appears to be between spatial and nonspatial thinking. But I will present reasons for dissatisfaction with any binary classification: we should aim instead to develop a much more discriminating taxonomy of kinds of mathematical thinking based on detailed cognitive research. (Giaquinto, 2007, p. 240)

In order to describe the division between geometric and algebraic thinking (or, better, between diagrammatic and linguistic representations), we can deploy the Wittgensteinian notion of *family resemblance*. Diagrammatic representations can differ widely, but we can identify features that we generally attribute to them. Then, given a particular instance of a diagram we can wonder what it has in common with other members of the same family.[53]

It is clear that CDs can be deployed to sustain Giaquinto's point. As Weber (2013, p. 64) notes, "images and inscriptions become rather indistinguishable in commutative diagrams." This means that according to him, diagrammatic and linguistic elements are integrated in the case of CDs.

[51] The same point has been made for knot diagrams in De Toffoli & Giardino (2014) and for categorical diagrams in sketch theory in Halimi (2012).

[52] On the other hand, Weyl (1932) famously distinguished sharply between these two different styles of reasoning in mathematics. I thank Tom Ryckman for pointing out to me this reference.

[53] Specifically, we can wonder how "prototypical" a particular representation is, as belonging to a certain kind. This is done with the aid of Rosch's (1999) application of the notion of "family resemblance" to psychology.

As a complement to the analysis of Giaquinto (2008, chap. 9), I add some elements to the list he developed to highlight the features which CDs share with prototypical diagrammatic representations:[54]

1. The represented content is displayed in a synthetic whole. (That is, all the relevant information is accessible at once in a unique display. In fact, one CD substitutes for many separate equations.)
2. Spatial relations in the diagrams correspond to algebraic relations (for example, two nodes are graphically connected if there is a homomorphism between them).
3. CDs exploit the two-dimensionality of the paper, instead of being formed by symbols in sequence.[55]
4. Different parts of the diagram have to be grouped together in order to use it efficiently. In the case of a diagram chase, experts have to identify different paths in the same diagrams. This corresponds to grouping together different subparts of the diagram, made by nodes and arrows (this is an instance of aspect shifting and visualizing paths).

This last point is well exemplified by Euclidean diagrams. In reasoning with Euclidean diagrams, a crucial operation is to interpret in different ways the same diagrammatic element. For example, in proving the possibility of constructing an equilateral triangle with side any given segment[56] we have to interpret the same segment as side of a triangle and as radius of a circle.[57]

CDs present also features which are typical of linguistic displays:

1. They are highly conventional.[58]
2. They have a syntax. This is formed by letters, indices, arrows, and a 'well-formedness' condition on the two-dimensional concatenation of these symbols. Moreover, they are recursive in the same way language is.
3. Isolating a specific part of the diagram, we get information relative to a specific bit of the algebraic structure.
4. Part of the information can be read-off sequentially. In fact, we can read single rows (or single columns), but then the information is not entirely sequential because we have vertical arrows as well.

Therefore, the *diagrammatic/linguistic* distinction, as well as the *geometric/algebraic* distinction, fails to capture the nature of CDs as well as the nature of other diagrams. This is not to deny the fact that these distinctions are meaningful in certain contexts, but to show that the dividing line is not sharp. Moreover, if we want to pursue a "more discriminating taxonomy of kinds of mathematical thinking," as Giaquinto envisages, then we have to abandon such oversimplified distinctions and introduce new ones, such as the one between intra-configurational and trans-configurational notations proposed by Macbeth.

---

[54] Giaquinto (2008, chap. 9) identifies the first two features of the list.

[55] This feature of diagrammatic displays is stressed by Krämer in (1988).

[56] This is Proposition I.1 of the *Elements*.

[57] See Macbeth (2012c, p. 69) for a detailed analysis of this example.

[58] Although conventionality is clearly a typical feature of linguistic displays, it has been argued that also all visual representations are "highly conventional," see for example Panofsky (1991) and Goodman (1976). Thanks to one of the referees for this note.

The extreme usefulness of CDs derives exactly from the combination of elements belonging to both sides of the line, as we saw in the previous section, when we described them in terms of their expressiveness, calculability, and transparency.

4.5. CDs in different mathematical contexts. Concerning the role of diagrams in different mathematical activities, traditionally the debate has focused almost exclusively on the role of diagrams in proofs.[59] In the following, according to this new trend in philosophy of mathematics, I will also take into consideration the roles diagrams can assume in phases of the mathematical enterprise other than proving, such as discovering and understanding. The latter is an important phase of the practice of mathematics, even if often its importance has been neglected.

Supporting the importance of evaluating the role of diagrams in this broader way, Carter (2010) presents a case study involving diagrams that are not part of proofs, but play nonetheless important roles. The diagrams she considers do not depict geometrical objects, but belong to free probability theory. In particular, she identifies two roles: (1) diagrams suggest definitions and proof strategies and (2) diagrams function as "frameworks" in parts of proofs.

In the case of homological algebra, CDs play the two roles identified by Carter: (1) a proof by diagram chase is inspired by the CDs and (2) if we consider CDs as maps depicting an algebraic situation, then they constitute the framework which allows us to understand and design the proof. Moreover, they are also the framework in which the statement of various results is expressible in a meaningful way.

CDs also play two other important roles: (3) They display algebraic content in a meaningful and mathematically tractable way. Without them, even the statement of a result such as the 5-lemma would be much more difficult to understand.[60] As we have seen, CDs organize content in such a way as to present the material in an accessible way, i.e., in a form experts can understand and use in order to draw correct inferences. (4) CDs allow for calculations by diagram chase; these calculations are an integral part of proofs involving CDs.

One difference with Carter's case is that CDs are mathematical objects in themselves that are essential not only to describe or represent some other mathematical object, but they are themselves the subject-matter.[61] This is similar to what Netz has argued for the case of diagrams in Greek mathematics:

> This shows that Greek mathematics relies upon diagrams in an essential, logical way. Without diagrams, objects lose their reference; so, obviously, assertions lose their truth-value. Ergo, part of the content is supplied by the diagram, and not solely by the text. The diagram is not just a pedagogic aid, it is a necessary, logical component. (Netz, 1998, p. 34)

The case of CDs is analogous: They are the subject-matter. That is why even stating a result without them would be difficult, let alone proving it. A similar phenomenon is

---

[59] See for instance Shin et al., (Fall 2013).

[60] For the FTHA, the statement is about a CD, but the representation is linear, by using a convention on the indices. Nevertheless, the two-dimensional diagram is essential not only to understand the proof, but also to clearly conceive the situation under investigation.

[61] It is important to keep in mind that a CD can be interpreted as a mathematical object or else as a concrete inscription, that is as a 'cognitive representation'. I am indebted to Marcus Giaquinto for helping me in clarifying this distinction.

observable in physics with Feynman diagrams, not only are they an inferential aid, but they make possible the exposition of content in a meaningful way.[62]

Feferman considers the proof of the FTHA and claims:

> While that proof [...] is written out fully in symbols, anyone who studies it can hardly deny that the diagram [...] is absolutely indispensable for understanding how it proceeds by "diagram chasing," i.e., the demonstration that the composition of maps along various paths from a given node to another in the diagram always gives the same result. This is completely typical of arguments in homological algebra, combinatorial topology and modern algebraic geometry. (Feferman, 2012, p. 379)

Feferman discusses the possibility of formalization of such diagrams, in order to address the worries concerning the presence of diagrams in proofs. It is clear that if proofs are conceived as a sequence of symbols, as it is traditionally the case, then diagrams find no place in them. Nevertheless, observing this case study and other cases[63] hints at the fact that the practice of mathematics does not follow such a strict definition of proof, and what counts as valid reasoning can contain diagrams as well. Moreover, the strict definition of proof as a sequence of symbols is essentially motivated by the desire of accepting as valid only "mechanically checkable proofs." In the case of CDs, this desideratum can be maintained, even admitting diagrams in proofs.

The fact that we can translate CDs into a sequence of symbols does not mean that their specific planar nature is not essential for mathematicians to prove results with them. If we consider a proof as a syntactical object made by symbols and independent of human understanding, then CDs find no place in a proof. But if we consider a proof what actual mathematicians call 'proof', then CDs are an essential part of many proofs in homological algebra.

Moreover, phrasing the problem of diagrams in proofs in terms of the possibility of eliminating them by translating the proof into a sequence of symbols misses the core of the issue. In fact, from the cognitive prospective, reformulating a proof analytically changes it completely. From my perspective, the fundamental problem is not whether a proof can contain diagrams or not, but whether it carries a stable mathematical knowledge. It is in this spirit that, in the case of Euclid, Manders (2008) analyzed the factors of the Euclidean practice that are responsible for the stability of the results of Euclidean geometry. Similarly, in the case of CDs we witness to a stable practice that is therefore also not at issue with respect to foundational worries.

§5. Conclusions.

By analyzing the uses of CDs, I hope to have conveyed how a mathematical notation can be effective by combining diagrammatic and linguistic elements. The simultaneous presence of these elements is what supports the specific form of calculation that expert perform on CDs: diagram chase. It is their hybrid essence that makes them such a perspicuous mathematical notation: Through their linguistic elements, they can represent general algebraic structures; through their configuration in two dimensions, they clearly represent the relations between these algebraic structures.

The analysis in the present article has been made possible by the new approaches proposed by the philosophy of mathematical practice. In fact, it is only by a close scrutiny of

---

[62] See Wüthrich (2010) for an introduction on Feynmann diagrams and an analysis of their origin.

[63] See for example the case of low-dimensional topology in De Toffoli & Giardino (2015).

the practice of mathematicians that we can appreciate results such as the one presented in this article. This attention to the practice sets the ground in which scholars can contribute to more traditional debates as well, in particular by enriching philosophical notions such as 'notation', 'proof', and 'justification' in mathematics. Therefore, the attention to the practice leads us to shape the epistemological inquiries in philosophy of mathematics differently and more broadly. With the case of CDs, I hope to have contributed to this more general aim.

A starting remark of this inquiry was that the role of notations is crucial in mathematics. A natural continuation of this research would be to enlarge the set of analyzed cases and include different types of notations, both in mathematics as well as in other areas of studies, such as natural or social sciences. In this direction, there are a number of studies on diagrams in natural science and in architecture and design,[64] but a comparative study is still missing.

Another further direction of inquiry is found from the observation that the point of view on mathematical notations presented here opens the door to experimental lines of inquiries. In fact, different mathematical notations can be approached also from the ground of cognitive science. In particular, it would be useful to identify what are the specific cognitive abilities triggered by different notations. For example, certain conventions in diagrammatic notations are better than others since they exploit our cognitive abilities such as vision and agency[65] and the transparency of a notation (as defined in this article) depends on different cognitive abilities.

The case of CDs would be particularly interesting also from this point of view because in it different cognitive abilities are triggered by their heterogeneous features. As I aimed to show, various representational advantages converge in a single CD.

§6. Acknowledgements. I wish to thank Valeria Giardino for great suggestions and unending conversations on these topics. Moreover, I wish to thank: Jessica Carter, Tom Donaldson, Michael Friedman, Marcus Giaquinto, David Hills, Paolo Mancosu, Kenneth Manders, Thomas Ryckman, John Sullivan, David Waszek, and the referees for their helpful comments. I presented a version of this paper in the *Proofs and Pictures* conference in Columbia, SC (March 2015) and in the *Third International Meeting of the Association for the Philosophy of Mathematical Practice* in Paris (November 2015); I am thankful for the useful feedback received in such occasions.

BIBLIOGRAPHY

Adams, C. C. (1994). *The Knot Book*. New York: W. H. Freeman.

Alexander, J. W. (1928). Topological invariants of knots and links. *Transactions of the American Mathematical Society*, **30**(2), 275–306.

Ammon, S. (2015). Einige Überlegungen zur generativen und instrumentellen Operativität von technischen Bildern. In Depner, H., editor. *Visuelle Philosophie*. Würzburg: Königshausen & Neumann.

Avigad, J., Dean, E., & Mumma, J. (2009). A formal system for Euclid's *Elements*. *The Review of Symbolic Logic*, **2**(04), 700–768.

Bredon, G. E. (1993). *Topology and Geometry*, Vol. 139. New York: Springer.

[64] See for example Perini (2005) for natural sciences and of Ammon (2015) for architecture.

[65] See for example the case of knot theory in De Toffoli & Giardino (2014).

Brown, J. R. (2008). *Philosophy of Mathematics: A Contemporary Introduction to the World of Proofs and Pictures*. New York and London: Routledge.

Carter, J. (2010). Diagrams and proofs in analysis. *International Studies in the Philosophy of Science*, **24**(1), 1–14.

Colyvan, M. (2012). *An Introduction to the Philosophy of Mathematics*. Cambridge: Cambridge University Press.

De Toffoli, S. & Giardino, V. (2014). Roles and forms of diagrams in knot theory. *Erkenntnis*, **79**(3), 829–842.

De Toffoli, S. & Giardino, V. (2015). *An Inquiry into the Practice of Proving in Low-dimensional Topology*. Boston Studies in the Philosophy and History of Science, Vol. 308, Chapter 15. Boston: Springer, pp. 315–336.

Dutilh-Novaes, C. (2012). *Formal Languages in Logic: A Philosophical and Cognitive Analysis*. Cambridge: Cambridge University Press.

Feferman, S. (2012). And so on...: reasoning with infinite diagrams. *Synthese*, **186**(1), 371–386.

Giaquinto, M. (1994). Epistemology of visual thinking in elementary real analysis. *The British Journal for the Philosophy of Science*, **45**(3), 789–813.

Giaquinto, M. (2007). *Visual Thinking in Mathematics*. Oxford University Press.

Giaquinto, M. (2008). Visualizing in mathematics. In Mancosu, P., editor. *The Philosophy of Mathematical Practice*. Oxford: Oxford University Press.

Giardino, V. & Greenberg, G. (2015). Introduction: Varieties of iconicity. *Review of Philosophy and Psychology*, **6**(1), 1–25.

Goodman, N. (1976). *Languages of Art* (second edition). Indianapolis: Hackett Publishing Company.

Grosholz, E. R. (2007). *Representation and Productive Ambiguity in Mathematics and the Sciences*. Oxford: Springer.

Halimi, B. (2012). Diagrams as sketches. *Synthese*, **186**(1), 387–409.

Ishiguro, H. (1990). *Leibniz's Philosophy of Logic and Language*. Cambridge University Press.

Kanizsa, G. (1980). *Grammatica del vedere: saggi su percezione e gestalt*. Il mulino.

Krämer, S. (1988). *Symbolische Maschinen: die Idee der Formalisierung in geschichtlichem Abriß*. Wissenschaftliche Buchgesellschaft.

Krämer, S. (2003). Writing, notational iconicity, calculus: On writing as a cultural technique. *Modern Language Notes*, **118**(3), 518–537.

Lang, S. (2002). *Algebra* (revised third edition). Springer.

Larkin, J. H. & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science*, **11**(1), 65–100.

Macbeth, D. (2012a). Diagrammatic reasoning in Frege's *Begriffsschrift*. *Synthese*, **186**(1), 289–314.

Macbeth, D. (2012b). Proof and understanding in mathematical practice. In Giardino, V., Moktefi, A., Mols, S., & Bendegem, J. P. V., editors. *From Practice to Results in Logic and Mathematics*. Philosophia Scientiae, Vol. 16. Kimé, pp. 29–54.

Macbeth, D. (2012c). Seeing how it goes: Paper-and-pencil reasoning in mathematical practice. *Philosophia Mathematica*, **20**(1), 58–85.

Mancosu, P. (2008). *The Philosophy of Mathematical Practice*. Oxford University Press.

Mancosu, P., Jørgensen, K. F., & Pedersen, S. A. (2005). *Visualization, Explanation and Reasoning Styles in Mathematics*, Vol. 327. Springer.

Manders, K. (2008). The Euclidean diagram. In Mancosu, P., editor. *The Philosophy of Mathematical Practice*. Oxford University Press, pp. 112–183.

Netz, R. (1998). Greek mathematical diagrams: Their use and their meaning. *For the Learning of Mathematics*, **18**(3), 33–39.

Panofsky, E. (1991). *Perspective as Symbolic Form* (revised edition). New York: Zone Books.

Perini, L. (2005). Visual representations and confirmation. *Philosophy of Science*, **72**(5), 913–926.

Rosch, E. (1999). Reclaiming concepts. *Journal of Consciousness Studies*, **6**(11–12), 11–12.

Schlimm, D. & Neth, H. (2008). Modeling ancient and modern arithmetic practices: Addition and multiplication with arabic and roman numerals In Sloutsky, V. Love B. & McRae K. (Editors), 30th Annual Conference of the Cognitive Science Society (pp. 2097–2102). Austin, TX: Cognitive Science Society.

Shin, S.-J. (2004). Heterogeneous reasoning and its logic. *The Bulletin of Symbolic Logic*, **10**(1), 86–106.

Shin, S.-J., Lemon, O., & Mumma, J. (Fall 2013). Diagrams. In Zalta, editor. *Stanford Encylcopedia for Philosophy*. Available at: http://plato.stanford.edu/archives/fall2013/entries/diagrams/.

Simmons, H. (2011). *An Introduction to Category Theory*. Cambridge University Press.

Weber, Z. (2013). Figures, formulae, and functors. In Shin, S.-J. and Moktefi, A., editors. *Visual Reasoning with Diagrams*. Basel: Birkhäuser, pp. 153–170.

Weyl (1995 (original in Gerlman 1932)). Topology and abstract algebra as two roads of mathematical comprehension. *The American Mathematical Monthly*, **102**(5), 453–460.

Winther, R. G. (2017). *When Maps Become the World*. Chicago: The University of Chicago Press.

Wüthrich, A. (2010). *The Genesis of Feynman Diagrams*, Vol. 26. Netherlands: Springer Science & Business Media.

DEPARTMENT OF PHILOSOPHY
STANFORD UNIVERSITY
450 SERRA MALL, STANFORD
CA 94305, USA
E-mail: silviadt@stanford.edu