Reasoning by Analogy in Mathematical Practice

Nicolò Cangiotti* and Francesco Nappo*

*Department of Mathematics, Politecnico di Milano

How do mathematicians make mathematical discoveries? Any historically accurate answer to this question cannot ignore the role of analogies. Explaining his efforts in the theory of fields, notable mathematician André Weil writes to his sister Simone (the famous philosopher): "Artin functions *in the abelian case* are *polynomials*, which one can express by saying that these fields furnish a *simplified* model of what happens in number fields; here, there is thus room to conjecture that the non-abelian Artin functions are still polynomials; *that is just what occupies me at the moment*" (in Krieger 2005:340). As Weil elaborates, this way of making conjectures about the features of an unfamiliar domain from those of a more well-understood analogue is a central part of a mathematician's methodological toolbox. As a matter of historical fact, "nothing is more fecund than these slightly adulterous relationships [between disparate mathematical domains]; and nothing gives greater pleasure to the connoisseur" (*ibidem*, 339).

From a methodological perspective, it seems natural to equate the role of analogy in mathematical research to that of a useful heuristics - a psychological aid for generating conjectures to be proved or disproved. However, a closer look reveals that the attribution of a heuristic role insufficiently captures what analogies mean for the working mathematician. As a matter of fact, an analogy with a well-understood mathematical domain is often the basis of considerable expectations about the truth of a conjecture. When the aim is to devise a proof, entire research programs are initiated and pursued that exploit analogies with mathematical problems already solved. These aspects of mathematical practice indicate an implicit commitment to the view that drawing an analogy with a familiar mathematical domain can be not only a useful heuristic, but is sometimes the source of *inductive support* for yet unproven conjectures, i.e., some (possibly small, yet non-negligible) additional ground for believing them.

We find a contemporary illustration of the inductive role of analogy in the history of the Poincaré conjecture. Formulated by Poincaré in 1904, this theorem of algebraic geometry states that every simply connected, closed 3-manifold (a generalization of Euclidean space) is homeomorphic to a 3-sphere (a higher-dimensional analogue of a sphere). While resisting several attempts at a proof, the conjecture was made plausible by an analogue theorem in the two-dimensional case, viz., that every simply connected, compact 2-dimensional surface without boundaries is homeomorphic to the 2-sphere (the 'ordinary' sphere). The proof-idea was later found with the help of another analogy, with the so-called 'Ricci flow with surgery'. In brief, Hamilton (1982) noted that one could perform operations on 3-manifolds that are mathematically analogous to the action of the

heat equation; he thus proposed that the effect of a Ricci flow on the former tends to a behavior analogous to the uniform behavior that the heat equation tends to under the same operation. Eventually, Perel'man (2003) closed the remaining gap in Hamilton's proof-idea, at a time when practically no more doubt remained about the conjecture's truth.

Our aim in this paper is to make progress on the 'descriptive' project to provide an account of the conditions under which an analogy in mathematics is considered capable of providing support to a mathematical conjecture. The conditions that we will put forward aim to apply to both analogical arguments used in support of a hypothesis and those used in support of some undemonstrated lemma which helps prove a hypothesis. We will defend our framework for distinguishing strong from weak analogical inferences in mathematics by reference to case-studies. One of the main promises of our account will be its potential to clarify how analogical reasoning can be used to extend mathematical properties known to hold in the finite case to infinite domains.

An important theme in the talk will be elaborating the divergences between the account proposed in our paper and Bartha's (2009) - arguably the most developed attempt at an account of plausible analogical inference in the current philosophical literature. As we will argue, Bartha is too quick to abandon the 'two-dimensional' framework for analogical reasoning in empirical sciences put forward by Hesse (1963). Contra Bartha's claim that "in mathematics… material analogy, in Hesse's sense, plays no role at all" (43), we will defend the necessity of restricting inductively significant analogical inferences in mathematics to those which rely upon *'material'* or *'pre-theoretic'* similarities in Hesse's sense. Furthermore, we will argue that, corresponding to Hesse's distinction between analogical arguments in science that project *causal relations* from source to target and those that project merely *statistical correlations*, there exists a parallel need to restrict inductively significant analogical inferences in mathematics to those that project some *robust mathematical connection* from source to target domain. Our case-studies aim to expose the dramatic divergences with Bartha's account in a clear and accessible manner.

As an illustrative example, below we will focus on the case of geometrical analogy that Bartha (2009) also analyzes. The source in this analogy is the geometry of triangles. By the 'median' of a triangle, we mean the segment that unites the midpoint of a side to the opposite vertex. The following theorem holds:

> *The three medians of a triangle intersect in a common point (a 'barycenter' or 'centroid').*

The question is whether this fact about triangles supports, by analogy, the conjecture:

> *The four medians of a tetrahedron intersect in a common point,*

where 'median' here refers to the segment that unites a face's centroid with the opposite vertex

On Bartha's (2009:110) "articulation model" of analogical reasoning, the strength of the inference from the two-dimensional to the three-dimensional case depends on the source proof

that one picks as what he calls the 'prior association' for the argument – the relation between the known and the predicted similarities. In one version, the source proof involves Ceva's theorem and is purely geometric. From the definition of median of a triangle ABC, we have:

$$\frac{AX}{XB} \cdot \frac{BY}{YC} \cdot \frac{CZ}{ZA} = 1$$

where X, Y, and Z are the three midpoints of the segments AB, BC, and CA respectively. It is a consequence of Ceva's theorem that segments AY, BZ, and CX are concurrent in a point. An alternative source proof for the argument employs analytic geometry and the Euclidean scaffold (i.e., a Cartesian coordinate system). The definition of a midpoint in a triangle is:

$$X = \frac{A+B}{2}; \qquad Y = \frac{A+C}{2}; \qquad Z = \frac{B+C}{2}.$$

By considering each median in its analytical form (e.g., $X \cdot t + (1 - t) \cdot C$ for median XC), we calculate that there is a point that lies on each median, corresponding to the value $t = \frac{2}{3}$.

On Bartha's view, the case of inferring properties of tetrahedra from analogous properties of triangles demonstrates an interesting form of *proof-sensitivity*: the analogical inference to the four medians of a tetrahedron intersecting in a common point is "not plausible" (2009:110), on Bartha's account, when we take the demonstration via Ceva's theorem as the source proof; it is only "plausible" (110) when we take the demonstration via analytic geometry as the source proof. In the latter case, we can rely on the observation that the centroids of a tetrahedron are:

$$X = \frac{A+B+C}{3}; \qquad Y = \frac{A+B+D}{3};$$

$$Z = \frac{B+C+D}{3}; \qquad W = \frac{A+C+D}{3}.$$

In this way, we can exhibit a resemblance with the definition of midpoints in the two-dimensional case. By contrast, the analogical inference via pure geometry allegedly suffers from the fact that "there is no clear three-dimensional analogue of Ceva's theorem" (110) and hence Bartha's "*No-Critical-Difference* condition fails" (110). The latter is his condition that, in order for an analogical argument to be regarded as plausible, "no explicit assumption in the source proof can correspond to something known to be false in the target" (2009:159).

As we will discuss in the talk, we find that Bartha's account gets the plausibility judgments about this case-study (and many others) wrong. The strength of the analogical inference from the two-dimensional to the three-dimensional case depends mainly, on our view, on a purely geometrical intuition: given that the 'median' in both the two-dimensional and the three-dimensional case is what unites the 'center of mass' (so to speak) of a figure or face to the opposite vertex, it is plausible that the medians in both the triangle and the tetrahedron case will concur in a point. Once we abandon the context of pure geometry and move to representing triangles and tetrahedra in a Cartesian coordinate system, the analogical inference arguably

weakens. First, the similarities upon which we rely become thinner. Indeed, we are left merely with a similarity in form between the expressions for midpoints and centroids, namely:

$$\frac{A+B}{2}, \quad \frac{B+C}{2}, \text{ etc. } \text{ resemble } \quad \frac{A+B+C}{3}, \quad \frac{A+B+D}{3}, \text{ etc.}$$

Second, and perhaps most importantly, abandoning the context of pure geometry results in a loss of generalizability of the vertical relations in the analogy. When considering the problem in purely geometrical terms, it immediately strikes us as a serious epistemic possibility that the passage from triangles to tetrahedra does not affect the geometrical connection between that the triangle's midpoints (their 'centers of mass'), medians, and barycenter in a triangle. Indeed, knowing the fact about triangles, it seems reasonable to expect that a geometrical connection of the same kind holds not just for three-dimensional Euclidean spaces, but also in spherical and hyperbolic geometry. By restricting ourselves to analytic geometry, instead, our trust in the possibility of generalizing connections that we know to hold in that restricted context is eroded. For instance, while in principle it is possible to obtain an analytical generalization of the theorem about triangles in higher dimensions or in alternative spaces, in each case the formalization is far from trivial and, most importantly, it does not arise naturally from the analytic background.

The lesson we draw from this case-study is the following: it is a mistake to recommend (as Bartha's account does) that the analogical inference from triangles to tetrahedra be evaluated within the context of analytic rather than pure geometry, simply because of some property of the *proof* of the theorem in the source domain. Indeed, once we shift to describing the problem in analytic terms, the similarities that the inference relies upon become thinner and the 'vertical relations' are arguably less suitable candidates for projection. On our view, the analogical inference is much stronger when we consider it within the context of pure geometry that is its natural home. In that case, we are able to more clearly focus on the geometrical reason for why two seemingly different elements, namely median of a triangle and median of a tetrahedron, are both appropriately named 'medians'; from this underlying intuition, we can plausibly conjecture that in both cases the medians are concurrent to a point. Hence, we will defend a different condition on plausibility than Bartha's No-Critical-Difference Condition, which goes as follows:

> *Relevance*: the vertical relation in the argument (i.e., the relation between the known and the predicted similarities) must be some robust mathematical connection; and it must be a serious possibility that a mathematical connection of the same kind as the source's also obtains between the known and the merely predicted properties of the target.

From considering further case-studies of analogy in mathematics, we will propose some further conditions in addition to *Relevance* (which we are not able to illustrate in this abstract; see Cangiotti and Nappo ms. for details). Our account turns out to be a generalization of the criteria for strong analogical reasoning in science put forward by Hesse in her groundbreaking (1963). Accordingly, our discussion can be considered a defense of the relevance of Hesse's notion of *material analogy* for capturing the practice of analogical inference in the domain of mathematics.

Bibliography

Cited in this abstract:

Bartha, P. 2009. *By Parallel Reasoning*. New York: Oxford University Press.

Cangiotti, N. and Nappo, F., ms. "Reasoning by Analogy in Mathematical Practice". Pre-print available at: http://philsci-archive.pitt.edu/20202/

Hamilton, R. S. 1982. "Three-manifolds with positive Ricci curvature". *Journal of Differential Geometry*, 17: 255-306.

Hesse, M. 1963. *Models and Analogies in Science*. Notre Dame: Notre Dame University Press.

Krieger, M. 2005. "A 1940 Letter of André Weil on Analogy in Mathematics". *Notices of the AMS*, 52(3): 334-341.

Perel'man, G. 2003. "Finite extinction time for the solutions to the Ricci flow on certain three-manifolds". Preprint on arXiv: https://arxiv.org/abs/math/0307245.

Other relevant works:

Corfield, D. 2008. *Towards a Philosophy of Real Mathematics*. Cambridge, UK: Cambridge University Press.

Franklin, J. 2013. "Non-Deductive Logic in Mathematics: The Probability of Conjectures". In Aberdein, A. and Dove, I. (eds.), *The Arguments of Mathematics*, New York: Springer, 11-29.

Goodman, N. 1954. *Fact, Fiction and Forecast*. Cambridge, MA: Harvard University Press.

Hacking, I. 2014. *Why Is There Philosophy of Mathematics at All?* Cambridge, UK: Cambridge University Press.

Hadamard, J. 1954. *The Psychology of Invention in the Mathematical Field*. New York: Dover.

Lakatos, I. 1976. *Proofs and Refutations*. Cambridge, UK: Cambridge University Press.

Poincaré, H. 1905 [1958]. *The Value of Science*. New York: Dover.