Topoi
https://doi.org/10.1007/s11245-018-9549-1

CrossMark

Manufacturing a Mathematical Group: A Study in Heuristics

Emiliano Ippoliti[1]

Abstract

I examine the way a relevant conceptual novelty in mathematics, that is, the notion of group, has been constructed in order to show the kinds of heuristic reasoning that enabled its manufacturing. To this end, I examine salient aspects of the works of Lagrange, Cauchy, Galois and Cayley (Sect. 2). In more detail, I examine the seminal idea resulting from Lagrange's heuristics and how Cauchy, Galois and Cayley develop it. This analysis shows us how new mathematical entities are generated, and also how what counts as a solution to a problem is shaped and changed. Finally, I argue that this case study shows us that we have to study inferential *micro-structures* (Sect. 3), that is, the ways similarities and regularities are sought, in order to understand how theoretical novelty is constructed and heuristic reasoning is put forward.

Keywords Heuristics · Hypotheses · Inference · Inferential micro-structures · Logic · Discovery

1 Introduction

In this paper I examine the construction of the concept of group in mathematics in order to discuss fundamental heuristic procedures. This concept and its history have been studied extensively (see e.g. Barnett 2010, 2017; Birkhoff 1937; Chakraborty and Chowdhury 2005; Chowdhury 1995; Kleiner 1986, 2007; Ronan 2006; Wussing 1984), so I will approach it from a heuristic viewpoint, that is, by focussing on the heuristics that gradually have led generations of mathematicians to its formation and refinement, and I will consider what we can learn from it. To this end, I will examine specific aspects of the works of Lagrange, and then I will look at how these seminal works have been developed by Cauchy, Galois and Cayley in order to produce a mature group theory. In more in detail I will examine the heuristic reasoning of Lagrange's work (Sect. 2.1), its development in the works on permutations by Cauchy, the introduction of the notion, and term, of groups provided by Galois, and then the first abstract treatment of group produced by Cayley (Sect. 2.2).

This analysis enables us to shed light on mathematical practice and to show how the manufacturing of the group concept displays paradigmatic features of the core of problem-solving, that is, the generation of a new hypothesis. To provide a bit more detail, I will examine how a new mathematical term, and concept, are introduced (i.e. the term group, *groupe* in French) and I will discuss if this case study agrees with major accounts of this issue, in particular I will look at Lakatos (1976); Grosholz (2007); Grosholz and Breger (2000); Cellucci (2013, 2017). I will also consider the notion of solution to a problem, or better what counts as a solution, and the way it changes.

Tellingly, this examination enables us also to shed light on the very first steps that make possible the formation of a heuristic procedure and then of a hypothesis, what I call 'inferential microstructures'. These microstructures, basically, are similarities and local regularities. I will show how these fundamental microstructures, that is, the ways similarities are sought and constructed, enable the formation of basic ampliative inferences. Therefore, I will argue (Sect. 3) that in order to account for scientific discovery and heuristic reasoning we have to examine its 'micro-structures', that is, to deepen our understanding of a different level of analysis of the inferential processes, the one that makes heuristic rules possible and applicable, as these rules presuppose them. Since these microstructures shape the construction of ampliative inferences, such as analogies, a better understanding of them would enable us to better understand the way pieces of information are introduced into the target of a problem and that the target did not contain at the beginning of the process.

Emiliano Ippoliti
emiliano.ippoliti@uniroma1.it

[1] Sapienza University of Rome, Rome, Italy

Journal : Large 11245	Article No : 9549	Pages : 9	MS Code : TOPO-D-18-00005	Dispatch : 17-2-2018

2 A Fine-Grained Analysis of the Manufacturing of Groups

The concept of group, and later of *abstract* group, was introduced and developed in order to solve a long-standing mathematical problem: to find an exact solution (not an approximate, that is, a numerical) to polynomials and polynomial equations[1] of a single variable *x*.

This problem can be dated back to ancient Greek mathematics, in particular geometry. In fact, several geometrical problems, such as ‘doubling the cube’ (Or Delian problem)[2], can be translated into modern algebraic symbolism—in this case $x^3 = 2$, namely a cubic equations. The search for a solution to those problems has been a powerful heuristic engine, as:

1. New mathematical entities were generated during the process;
2. What counts as a solution to a problem was reshaped;
3. The process built upon an initial ambiguous use of new terms and concepts, like the word ‘*group*’.[3]

To provide a bit more detail, as concerns (1), new curves like conic sections (i.e. parabolas, hyperbolas and ellipses) were introduced and studied by the Greek mathematicians in order to construct the line segments that could solve the geometrical problems, as problems like the Delian one is not solvable by using compass and straightedge, and many ingenious solutions where found by Greek mathematicians (e.g. Archytas, Menaechmus, Nicomedes, Diocles) looking for intersections of curves in space.

As concerns (2), the idea of what is an admissible solution to this problem changed over time. While for ancient Greek and, later, Islamic mathematicians, a negative number cannot be considered a solution (i.e. coefficients or roots of a polynomial), the introduction of complex numbers removed this constraint, reshaping the problem and its possible solutions.[4]

As concerns (3), the introduction of a new concept often goes through the search for a new term to express it and its novel content. A stock example is the Latin term *acies* used by Euler in his study of polyhedron formulas as noted by Lakatos (1976, p. 6). In effect a key to Euler’s result was just the introduction of the concepts of vertex and edge: he pointed out that, besides the number of faces, the number of points and lines on the surface of the polyhedron determines its topological character. To substantiate this novelty, Euler introduced the term ‘acies’ (edge) instead of the old *latus* (side), since latus was a polygonal concept while he wanted a polyhedral one.[5] Euler’s introduction of a new property and term was crucial in the search for (topological) invariant for a polyhedron. A new term, especially at the beginning of its usage, can be ambiguous since it can be imported from other domains and the meaning and content of the source affect the target.

2.1 Lagrange’s Heuristics

In order to solve the problem of finding an exact solution to polynomial equations, the Italian mathematician Giuseppe Lagrangia (Joseph Lagrange) produced the hypothesis of a relation between permutations and the solution of equations by radicals. This hypothesis is the keystone of Galois’s theory and, in general, of the construction of the mathematical treatment of groups.

How did he produce this hypothesis? In order to answer this question we need to examine the heuristic procedures employed by Lagrange, which will reveal two crucial aspects of heuristic procedures:

1. The construction of ‘inferential microstructures’ and then primitive heuristics over them;
2. What counts a solution to a problem.

Lagrange’s first heuristic move was to look for similarities, that is, common features of the known solution methods for specific polynomials—quadratics, cubics, and quartics[6]:

[1] That is a polynomial that has been set equal to zero—e.g. $x^2 + 7x + 5 = 0$.

[2] This problem requires, once provided the edge of a cube, finding the edge of a second cube whose volume is double that of the first (by using only the tools of a compass and straightedge), in modern algebraic symbolism$x = \sqrt[3]{2}$.

[3] Another interesting issue that characterizes this process (which we do not have room to treat here) is the heuristic role of new symbolisms. Partial solutions discovered in the sixteenth century increasingly made use of symbolism in a way that made possible theoretical questions to be posed and answered by mathematicians. For example, the refinement of the algebraic symbolism also allowed questions about the relation of roots and factors to be formulated and pursued (see Barnett 2010, p. 2 on this point).

[4] To provide a bi more detail, first, negative solutions were accepted, thus extending the domain of integers to the negative axis, so that, e.g., every first-degree equation in the normal form $ax + b = 0$ (even with positive coefficients a, b) has a solution. Then, by observing that second-degree equations, like e.g. $x^2 + 1 = 0$ have no solutions even in the extended realm of negative numbers, a further extension to complex numbers was introduced.

[5] He retained the term *angulus solidus* (solid angle) for his point-like vertices.

[6] For instance the methods of Cardano, Tschirnhaus, Euler, Bezout for cubics; and the methods of Cardano, Descartes, Tschirnhaus, Euler, Bezout for quartics.

"I examined and compared the principal methods known for solving algebraic equations" (Lagrange 1808, p. 1884, *my translation*). This analysis of several features of known solutions showed one common property: "I found that the methods all reduced, in the final analysis, to the use of a secondary equation called the resolvent" (*Ibid.*). That is, he found the existence of a secondary equation (the resolvent) whose roots would allow us to find the roots of the originally given equation. In other words, the solutions of the 'reduced' (or 'resolvent') equation are functions of the roots of the equation to be solved.

For instance, given a cubic equation (that is, $ax^3+bx^2+cx^3+d$) with three roots x_1, x_2, x_3, there is a reduced equation whose roots take values $y=\frac{1}{3}x_1+\alpha^2x_2+\alpha x_3$ and where $\alpha^3=1$, and $\alpha\neq 1$. Thus, Lagrange noticed that as y^3 takes just two values as the three roots are permuted, it satisfies a quadratic equation, which is just the *resolvent* of the cubic. A similar property can be found for quartics ($ax^4+bx^3+cx^2+dx+e$). In fact, given a quartic with four roots x_1, x_2, x_3, x_4 there is a resolvent whose roots are values, for instance, of $y=\frac{1}{2}(x_1x_2+x_3x_4)$. Thus, Lagrange noticed that as y takes just three values as the four roots are permuted, it satisfies a cubic equation.

In order to find these similarities he needed to manipulate the several sources (in this case, known methods of solution) in specific ways. In more detail, he sought for several ways of representing the 'resolvents' and tried to assimilate them. Thus, constructing a new representation is the first, decisive heuristic move. This change of representation allows us to find similarities, and then to build a inferential 'microstructure', that is the first tentative content of a inference, which can be a basis for primitive heuristic rules such as analogies and inductions. In effect, Lagrange's next heuristic move, after the similarities were found, was to try to generalize the suggested solution over all the possible cases of polynomial equations (i.e. a induction). In this way, he also reshapes the problem, as he aims at a unifying approach, and not simply at solving specific, small classes, of polynomial equations.

The starting point of his analysis is an algebraic property that is well-known at that time, namely the relationship between the roots and the coefficients of an equation. Here I will recall only the fundamental steps[7] of the analysis that led him to the formulation of his main hypothesis, that is, the idea of permuting roots.

1. Starting from relationship between the roots and the coefficients of an equation, the next step is to show (Lagrange 1808) that given the general equation $x^m - Ax^{m-1}+Bx^{m-2}-Cx^{m-3}+\ldots=0$, with its m roots x_1, x_2; $x_3,\ldots, x_m$, for the coefficients A, B, C, we have that:

   $A = x_1 + x_2 + x_3 + \ldots + x_m;$

   $B = x_1x_2 + x_1x_3 + \ldots + x_2x_3 + \ldots\ ;$

   $C = x_1x_2x_3 + \ldots;$

   Then, Lagrange notices that in the coefficients A, B, C,…, as the roots x_1, x_2,x_3, …,x_m are permuted, the formal value of the expression does not change[8]: the expressions resulting from this particular permutation of the given roots are formally equivalent to the original expressions. This occurs with every possible permutation of the roots for cubics and quartics.

2. In order to find a resolvent, that is, an equation whose solution would enable us to find an algebraic solution of the original equation, Lagrange (1808) argues that we can use the following economic method: *first find the form of all the roots of the sought equation, then compose this equation by means of its roots.*
3. He then shows how the permutation of the roots of the original equation in the formula for the resolvent's roots produces as many as *m!* resolvent roots for $m=3$ (i.e. cubes). Of course, since $m! > m$ for $m>2$, Lagrange's conclusion that an equation of degree *m* has a resolvent of degree *m!* not only offer no solution but also fails even to make progress toward one.
4. In order to overcome this difficulty, he argues the we can reduce the resolvent's degree by looking at some 'invariant'[9] in its roots and that the key to reducing the degree of the resolvent in the general case will be to consider the form of these roots, $t=x_1+ax_2+a^2x_3+a^3x_4+\ldots+a^{m-1}x_m$, and the effect of permutations on this form.

Unfortunately, this line of argument faces several difficulties even with relatively small values of *m*. In effect, even if a resolvent for a quintic equation can be reduced from degree 120 (5!) to degree 24 (4!), 24 is still considerably larger than the original equation's degree of 5. Similarly, for quartics, the initial resolvent degree of 4! = 24 that can be reduced only to 3! = 6.

As we know nowadays, these difficulties cannot be overcome in this way. Nonetheless, Lagrange's heuristic strategy paved the way to the construction of a novel approach and

[7] For a detailed analysis see in particular Wussing 1984; Barnett 2010, 2017.

[8] For instance, for $m=2$, by exchanging x_1 with x_2 we have that $A=x_2+x_1$ and B x_2x_1. Both are equal to the original expressions, where $A=x_1+x_2$ and $B=x_1x_2$.

[9] The invariance is expressed by Lagrange (1770) by saying that the resolvent "does not change" when *t* is replaced in a specific way.

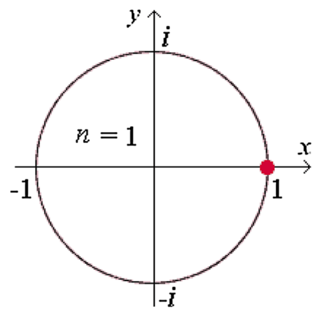

Fig. 1 Roots of unity on a complex coordinate system

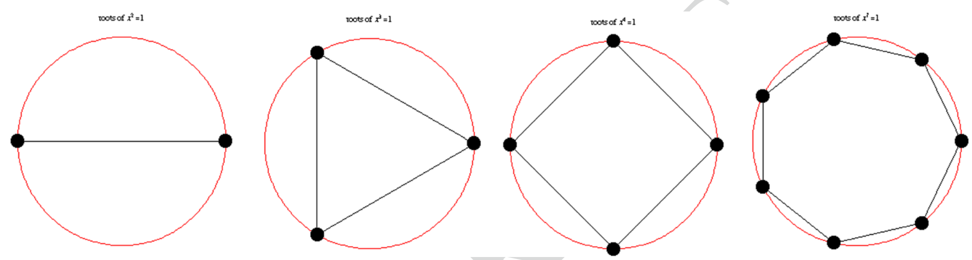

Fig. 2 Inscription of several polygons within the unit circle

concept in mathematics. Before examining it, it is worth noting that another interesting aspect of Lagrange's heuristics is the use of a trigonometric representation as an attempt of determining the roots of polynomials. This heuristic move, which produces another inferential microstructure, namely an assimilation of roots and points on a unit circle, is particularly important as it raises the issue of the solution to a problem, or better what can count as a solution to a problem. In effect, the use of a different representation raises the question of whether, and when, a specific representation is legitimate or admissible. The ways by which we set up a problem and represent it embed an idea of its possible solution, since these ways put specific constraints on the problem-space, that is the moves, the operators, and the intermediate states of our problem.

The geometric representation of roots, that is, roots of unity as points on the unit circle, was known in Lagrange's era and he employed it in order to find the total number and type (i.e. real or complex) of m^{th} roots of unity (Lagrange 1770). To provide a bit more detail, given a unit circle (i.e. with radius set equal to 1) on the complex coordinate system (see Fig. 1), the roots of unity, e.g. cube roots, fifth roots, etc., can be found by dividing the circle respectively into three, five, etc., equal parts.[10]

Therefore, the task of finding roots of unity is reduced to the inscription of polygons within the unit circle—a triangle for cube roots, a pentagon for fifth roots, etc. (see Fig. 2), where the roots are points in the circle.[11]

This representation is controversial. As Lagrange himself noticed, equations like $x^m - A = 0$, or more simply $x^m — 1 = 0$ (i.e. $x^m = 1$), are "always solvable by trigonometric tables in a manner that allows one to approximate the roots as closely as desired, by employing the known formula

$$x = \cos\frac{k}{m}360° + \sin\frac{k}{m}360°\sqrt{-1}$$

and letting $k = 1, 2, 3, \ldots, m$" (Lagrange 1770, p. 168). This formula is obtained from the polar, or trigonometric, form of complex numbers, which is one of the three forms by which we can represent a complex number. Unfortunately this approach does not work in general: a trigonometric function can be assimilated to an algebraic solution of polynomials only if the specific trigonometric values can be expressed in an *algebraic* form.[12]

This holds not only for this specific case: what counts as a solution to polynomials has been set up differently in different times. To illustrate this, let us think of the well-known example of the algebraic solution provided by the mathematician Omar Khayyam. He explicitly solved cubic equations by intersecting appropriate conic sections, that is, in a geometrical way, but not counting negative numbers as coefficients or roots of equations, since negative numbers were not allowed at that time. Since negative numbers were

[10] The roots of unity are all found on the unit circle, as they must have modulus equal to one. Now, every point on the unit circle can be identified by an arrow starting at the origin of coordinates and pointing at it. This arrow creates an angle alpha with the real axis (called *argument* of the complex number). Each multiplication of a root by itself rotates its arrow counterclockwise by the same angle alpha. So the *m*-roots of unity are such that after rotating *m* times by alpha one gets back to the point $x = 1$, i.e., m alpha = 360°, which has

Footnote 10 (continued)

exactly *m* distinct solutions $\text{alpha}_k = (k/m)\ 360°$, with $k = 0, \ldots, m-1$. Increasing *k* we aobtain solutions that are equivalent, e.g. $k = m$ gives alpha = 360° which is the same arrow as alpha = 0, $k = m + 1$ gives alpha = 360° + 360°/m, which is the same arrow as alpha = 360°/m, and so on.

[11] The roots are the points on the circle: their values have a real part and an imaginary part, and are measured respectively along the horizontal and vertical coordinates.

[12] By 'algebraic form' we mean that the roots of a given equation can be determined from its coefficients by means of a finite number of steps that involve only elementary arithmetic operations (+; −, ×, ÷), and the extraction of roots.

not an admissible solution, the search for a solution was *constrained* by this piece of knowledge and, as a consequence, the construction of the problem-space was affected by it. Of course the introduction and acceptance of complex numbers changed this: by allowing different kinds of solution the problem-space was shaped in a different manner.

This raises an important point about the notion of solvability, which can be set up in ways that can change over time. This change basically depends on the way we represent, and also *not* represent, a problem. Since a representation is built on inferential microstructures, that is similarities and local regularities, this means that the shape and boundaries of the problem-space can be modified, that is, extended or limited, by the production of specific inferential microstructures. Euler's solution to Basel problem is a stock example in this sense: when he provisionally admits as legitimate the representation of power series as trigonometric functions, this removes implicit constraints on the problems and changes the construction of the problem-space (see Ippoliti 2008).

Thus, we are now in a position to extract interesting features from Lagrange's heuristic strategy. In sum, Lagrange's heuristics can be broken down into the four following steps:

(s1) Change the ways of expressing the target of a problem, that is, of some of its features.

(s2) Look for similarities between ways of expressing the target of a problem or with other pieces of knowledge—that is, generate inferential microstructures.

(s3) Use these microstructures to build an analogy (a primitive heuristics, see Ippoliti 2018; Ippoliti-Cellucci 2016) and get a tentative solution (i.e. permutation of roots).

(s4) Use this tentative solution as a basis for a generalization (i.e. induction, that is, another primitive heuristics) and, in that case, if needed, refine the hypothesis in order to apply it to all the cases.

The outcome of this heuristic chain is the very tentative hypothesis that the solution of polynomial equations is a function of *permutations* of radicals. Now, as we know retrospectively, despite his success with polynomials of degree 3 and 4, Lagrange was not able to achieve a similar result for polynomials of higher degree. But this was not his fault: Abel's famous negative result (Abel 1824) showed that a 'quintic' formula for the general fifth degree polynomial is impossible to find and the same holds for equations of higher degree. Nevertheless, "Lagrange's introduction of permutations into the picture was the first significant step forward in the study of algebraic solvability in centuries" (Barnett 2017, p. 23).

In effect, the inferential microstructure introduced by Lagrange has borne fruit. First, Abel's proof is based just on the concept of permutation introduced by Lagrange. Second, the attempt of classifying the equations by means of solvability by radicals enabled Evariste Galois not only to introduce the term group ('*groupe*' in French), but also to construct the so-called 'group of permutations', which shape Galois theory. In turn, Galois' work on permutation groups provided the basis, and a specific instance, for a more general group concept as the one developed by Cayley (1854).

2.2 The Developments of Cauchy, Galois and Cayley

As often happens in history of mathematics, a failure not only is not fatal[13], but it can also provide a basis for building new knowledge. In effect, even if the heuristic chain s1–s4 built by Lagrange did not succeed, part of it, namely the inferential microstructure that moulded it (the idea of *permutations* of radicals), was used in different ways in mathematical problem-solving.

Cauchy explicitly developed Lagrange's microstructure and in a sense he made it an autonomous subject (see Cauchy 1815): in his works he does not deal with polynomial equations, but puts forward a systematic treatment of the algebraic properties of permutations.

In more detail, Cauchy applies Lagrange's idea (i.e. the number of distinguishable forms that result from permuting the variables in an expression is potential tool in studying algebraic solvability) more generally to any function of n variables, not simply to formulas for the roots of a resolvent equation. Basically, he tries to solve the following problem: "for a given number n of variables, what can be said about the possible number of distinct forms which a function of n variables can produce under permutations of those variables?" (Barnett 2010, p. 25).

In order to solve it, Cauchy introduces a new and better notation[14] for permutations, similar to a table of function values.[15] Endowed with such a new symbolism, he can introduce and, later, better study, the notion of composition of permutations and its properties (see Cauchy 1845). This opens the way to the construction and study of a *system of permutations*, today known as a *permutation group*.

[13] On this point see for instance (Gillies 1995; Ippoliti 2016). A recent interesting account is the one described by Fisch (2017), who deals with George Peacock's struggles that also ultimately failed but that, again, proved heuristically crucial in helping others to develop a new branch of mathematics.

[14] I won't discuss here another very interesting, crucial issue emerging from this case study, that is, the heuristic role of notation (e.g. notation allows generality, permits a useful polysemy, makes calculation possible, etc.), and in particular the way it allows problems to be *solved* and *posed*.

[15] Permutation here denotes the change of one *arrangement* (say *abc*) to another arrangement (e.g. *acb*)—in modern terms, we would say that it is a function that maps one set of entities (in this case, the letters *a, b, c*) onto the same set in a one-to-one manner.

Journal : Large 11245 | Article No : 9549 | Pages : 9 | MS Code : TOPO-D-18-00005 | Dispatch : 17-2-2018

Thus, while Lagrange introduces the concept of a permutation to deal with the algebraic solution of polynomial equations, Cauchy treats permutations as algebraic objects in their own right, and, in turn, this enables Galois and, later, Cayley to use permutations to construct respectively the notion of group and its abstract version.

Galois (1831) is the first one to use the word *group* in algebra. At the beginning, this term has an informal and ambiguous meaning: *group* initially occurs in the notion of *group of permutations* (*groupe de permutations*). In this context, *group* means *collection* or *set* in the loose sense of the word and Galois does not specify it or define it better. In more detail, Galois' idea is to consider a set and transformations from a set to itself that are invertible, so in this sense he is following Lagrange. On the other hand, Galois extends the heuristic reasoning of Lagrange and Cauchy by looking for the *internal structure* of a group of permutations as a way of controlling the solvability of the equation. Of course, Galois deals with groups of permutations, and not abstract groups, nonetheless, he looks at certain structural aspects of these groups that are isomorphism-invariant in our, modern, sense and that are independent of the particular action. Galois' theory applies to the roots of any polynomial, not just the to the cases where the roots are the variables and precisely in this sense he goes beyond Lagrange.

The permutation group is one of the explicit starting point of Cayley's theory of group—even if he makes it clear that he considers other objects and operations (e.g. quaternion imaginaries and their multiplication): "the idea of a group as applied to permutations or substitutions is due to Galois, and the introduction of it may be considered as marking an epoch in the progress of the theory of algebraic equations" (Cayley 1854, p. 124).

Another phenomenon that shaped the construction of the abstract notion of group in Cayley's works is not strictly mathematical in kind, as it comes from physics, namely geometrical optics (Cayley 1857). In particular, "Cayley's unexpected discovery of a non-abelian group of order 6 in the practical context of geometrical optics, served as the trigger for generalizing the group concept" (Chakraborty and Chowdhury 2005, p. 278). In more detail, this advancement was suggested by "Cayley's discovery of the six transformations that leave the equation of the secondary caustic unchanged, and his realization that these transformations form a group under the composition of mappings" (Chakraborty and Chowdhury 2005, p. 277). A caustic is a curve related to the reflection (or refraction) of light of a surface in the study of optics and this second, concrete, instance of a non-abelian group of order six suggested a generalization of the group concept beyond that of permutation group.

This shows how the assimilation of entities belonging to different domains was essential for the construction of the notion of group.

Cayley provides us with the first attempt of defining the modern, abstract,[16] concept of group. In particular, he writes: "a set of symbols 1, α, β,. . ., all of them different, and such that the product of any two of them (no matter in what order), or the product of any one of them into itself, belongs to the set, is said to be a group" (Cayley 1854, p. 124). In this sense, he produces "a remarkable conceptual novelty in mathematics, the birth of one of the most essential concepts of modern mathematics" (Pengelley 2005, p. 7). In effect, Cayley treats 1, α, β,. .. as completely abstract symbols—not simply permutations, quaternions, invertible matrices under multiplication, Gauss' quadratic forms, elliptic functions nor any other concrete instance of an operation. Consequently, any theorem that can be deduced from the properties stated in this definition will necessarily apply to every particular system that satisfies those properties.

Provided with this new explicit concept, Cayley is in a position to advance knowledge and problem-solving in algebra. For instance, the core of his work (1854) is the classification of groups of finite order according to their form. He shows that all groups of prime order p have the same form as the cyclic group, thus providing a classification for groups of prime order p, namely, that there is essentially only one group of order p for any given prime p. Moreover, he determines all the groups of orders 4 and 6, showing that there are exactly two of each.

In this sense the notion of group enables a classification of mathematical entities. More generally, the result obtained by Cayley shows the heuristic role of a new concept.

First, it provides a generalisation and unification of properties and entities. It makes a domain more compact by giving a unified treatment of previously heterogeneous mathematical entities. As a consequence, it boosts mathematical problem-solving, since a more general concept can be used to approach distinct entities, and local results and techniques can be transferred from one kind of entity to another kind—for instance from elliptic functions to invertible matrices. Thus it also provides a way of classifying entities and, accordingly, of seeking and establishing relations among them.

Moreover, it opens lines of research. For instance, the context of Cayley's work seemed to suggest the he was thinking of finite groups and so it was not clear if his results could hold also for infinite groups. This line of research was fruitful, as, in effect, later a formal proof of it was produced also for infinite groups.

[16] Cayley did not coin the term and notion of *abstract group*, as it emerged and became explicit later (see Chakraborty and Chowdhury 2005; Pengelley 2005).

Journal : Large 11245	Article No : 9549	Pages : 9	MS Code : TOPO-D-18-00005	Dispatch : 17-2-2018

Last but not least, it enables an abstraction that may end up with an axiomatization, as happened just with the construction of the abstract theory of groups.

3 The Role of Inferential Microstructures

The inferential process that led generations of mathematics, as we have seen, to construct the group concept required an assimilation[17] of different entities, i.e. invertible matrices, permutations, Gauss' quadratic forms, quaternions, various kinds of elliptic functions and caustics. This assimilation is based on an integration of pieces of knowledge coming from different fields and it shows how a conceptual novelty or change takes place in mathematics. Not only this assimilation generates a new object—the notion of group—but it also brings to light new features of already known objects by means of gradual steps that remodel these entities by adding pieces of information, that is, properties and relations, to them. In effect, the notions of permutation, matrix, or elliptic function are different from what they were before the construction of the notion of group: their properties, relation and inter-relation have changed, and moulded, in a new way and can serve new purposes in mathematical problem-solving.

Moreover, this example shows us, on one hand, that Lakatos' account of conceptual novelty and change is defective, and, on the other, that the more recent accounts (in particular see Grozholz, Cellucci) make better sense of the process.

Lakatos's seminal analysis of mathematical conceptual novelty (Lakatos 1976) focuses on proof-generated concepts, that is, ones based on the 'dialectics' between proofs and refutations. On one hand, Lakatos' analysis really grasps a way mathematical knowledge is advanced. On the other hand, such an analysis cannot account for the inferential, informal and rational work before the proof or the refutation, which is essential to understand how a mathematical novelty is introduced and, eventually, a theorem proved. In effect, Lakatos account cannot fully explain how the concept of group is introduced since it requires Lagrange's heuristics, and this involves the idea of studying the action of a structure on itself, which is not a part of a proof or a refutation. Moreover, Lakatos cannot fully explain the developments of Cauchy, Galois and Cayley, which use part of Lagrange's heuristics (s1–s4) and do not start from a well-established result, namely a proof (a theorem) or refutation.

Instead, this case study fits well with Grosholz's and Cellucci's accounts of mathematical novelty, as these accounts shed light on the inferential, informal and rational work before a proof or a refutation can be found.

On one hand, Emily Grosholz has showed that the employment of "modes of representation is typical of reasoning in mathematics" (Grosholz 2007, p. 4), and that the key to solving problems is not to "eliminate modes of representation", but to "multiply and juxtapose them" since "this often creates [...] productive ambiguity" (*Ibid*., xii). The construction of the mathematical group concept, as we have seen, required the integration of several representations, for instance a geometric and algebraic one. The formation of a new representation is a step-by-step process, which highlights certain features of the entity and neglects others. This construction, when successful, creates an information surplus that enables the solution of the problem. Moreover, as the case of the group concepts shows, the ambiguity of the term employed can play a heuristic role, since the multifaced aspect of an ambiguous concept, like that of 'group', enables the transfer of pieces of knowledge from one representation to another one. Of course, such a concept can be defined more precisely later and can have a more specific content, expressed even in an axiomatic fashion. So, the construction of a new representation might end up with a result that formalizes a kind of isomorphism or reduction between mathematical entities.

The fruitfulness of this process does not stop here. As a new representation is constructed and a result from it is proved in a given domain, it can be employed to deal with problems in other parts of mathematical knowledge, and so on, in a virtuous circle. That is the case of the concept of group and its abstract version, which is employed to solve problems in other mathematical domains and even in empirical scientific fields, like quantum mechanics.

On the other hand, Cellucci's account of conceptual novelty focuses on the role of ampliative inferences in the formation of a hypothesis to solve a problem: "hypotheses are obtained by non-deductive rules, rather than by deductive rules" since "non-deductive rules are ampliative, namely, the conclusion is not contained in the premisses" (Cellucci 2017, p. 154). In effect, the construction of the group concept required analogies in order to be put forward and be successful. The problem that triggered the formation of the group concept was changed and then solved with several heuristic reasoning: Lagrange's seminal heuristics, which produced the hypothesis that solution of polynomial equations is a function of permutations of radicals, was later refined and adjusted by Cauchy, Galois and Cayley.

What my analysis of the genesis and development of the notion of group shows more than these accounts is the fundamental role played by inferential micro-structures, that is, the search for similarities and local regularities. In effect, these micro-structures provide the building blocks that erect and shape the content of the conclusion of a ampliative inference.

17FL01 [17] See Thomas (2011) on the role of assimilation in mathematics.

Journal : Large 11245	Article No : 9549	Pages : 9	MS Code : TOPO-D-18-00005	Dispatch : 17-2-2018

The term 'microstructures' aims at highlighting the fact that they are *proto-inferences*: they provide a specific content for an ampliative inference, but they lack a sufficient level of generality, i.e. they do not allow obtaining a conclusion that is general in kind. Moreover, their content is not structured, that is, there is no hierarchical relation with other concept, as they basically are a way of searching for associations or correlations to refine.

Bottom line: they occur at a 'micro' level, that is, one that precedes the construction of a proper ampliative inference, and they provide the first, tentative, content of what later can be better structured, extended or refined as a more robust hypothesis.

The study of microstructures sheds light on the way pieces of information are introduced into the target of a problem and that the target did not contain at the beginning of the process. This information, which can take the form of new stipulations, functions, entities, not contained in the target at the beginning of the process of assimilation that shapes the construction of inferential microstructures. Since micro-structures are constructed by looking for properties of entities to assimilate, as we have seen with Lagrange's heuristics, they provide the basis for the construction both of new representations and of heuristic rules, such as analogies. Once an analogy is created, it can evolve in several ways: it might become a partial isomorphism, a kind of reduction, or even an identity.

Now, in order to build similarities, we need comparable aspects, and this in turn requires a certain viewpoint to compare things. This means that the way we represent things is essential to find possible similarities upon which to build plausible heuristic inferences. It is worth noting that the choice of a viewpoint is not arbitrary, as it is suggested by the features of the problem that we are trying to solve. It follows that first step in building a heuristic line of argument is the manipulation of the entities involved in the problems, a change of their representation, that is, adapting or changing them to serve a given purpose or advantage, even if it requires to neglect several features of the original entity.

The search for similarities in problems at the frontier of knowledge most of the time cannot be approached with metric or probabilistic measures of similarities. A numerical evaluation of similarities between entities, or of relevance of similarity, is hard to realize: theories of similarities, theories of relevance, or theories of typicality do no offer a cogent basis for analogical inference and heuristic reasoning (see Ippoliti 2006). The similarities, and the tentative content that they provide, have to be judged and weighed on a qualitative, case-by-case basis, but this does not limit the role of micro-structures.

The inferential microstructure, by looking for entities to assimilate, might lead us to novel entities, just like mathematical group, or even a new theory—algebraic topology is a paramount example in this sense (see Ippoliti 2016). The assimilation requires an integration of pieces of information belonging to several domains. This means that a new, emerging object is both unitary and multi-faced, and each 'face' can be used to solve a specific problem o sub-problem: the very same entity embeds different contents, the ones coming from different domains merged in it. As these faces, or pieces of information, contribute to the construction of these entities (are part of it), they ease the production of new representations, which in turn enable the manufacturing of new analogies and ampliative inferences. Similarities offer a basis for new representations and ampliative inferences, and these, in turn, offer a basis for properties or entities to be included in formal proofs.

So from similarities to theorems, the manufacturing of mathematical groups shows us the rational and inferential path that builds a novel concept, a path that starts with inferential microstructures that are revised and refined in order to pose, change and solve problems, and ends up with a new theorem or theory, even an axiomatic one such as abstract group theory.

Acknowledgements I would like to thank the two anonymous referees for the valuable advice on the both the philosophical and the mathematical parts of my paper.

References

Abel NH (1824) Mémoire sur les équations algébriques où on démontre l'impossibilité de la résolution de l'équation générale du cinquième degré. Groendahl, Oslo

Barnett JH (2010). Abstract awakenings in algebra: Early group theory in the works of Lagrange, Cauchy, and Cayley. draft

Barnett JH (2017). The roots of early group theory in the works of lagrange. Abstract Algebra 2. http://digitalcommons.ursinus.edu/triumphs_abstract/2

Birkhoff G (1937) Galois and group theory. Osiris 3:260–268

Cauchy A (1815) Memoire sur le nombre des valeurs qu'une fonction peut acquerir, lorsqu'on y permute de toutes les manieres possibles les quantites qu'elle renferme. J de l'Ecole Polytechnique 17(10):1–28

Cauchy A (1845) Mémoire sur les arrangements que l'on peut former avec des lettres données et sur les permutations ou substitutions à l'aide desquelles on passe d'un arrangement à un autre. Exercises d'analyse et de physique mathématique 3:151–252

Cayley A (1854) On the theory of groups as depending on the symbolic equation $\theta^n = 1$. Phil Mag 7(42):40–47

Cayley A (1857) A memoir upon caustics. Phil Trans R Soc Lond 147:273–312

Cellucci C (2013) Rethinking logic. Springer, Dordrecht

Cellucci C (2017) Rethinking knowledge. Springer, Dordrecht

Chakraborty S, Chowdhury MA (2005) Cayley and the abstract group concept. Math Mag 78(4):269–282

Chowdhury F, [1995]: Arthur Cayley and the theory of groups. J Math Math Sci 10:21–31

Fisch M (2017) Creatively undecided: toward a history and philosophy of scientific agency. University of Chicago Press, Chicago

Journal : **Large 11245**	Article No : **9549**	Pages : **9**	MS Code : **TOPO-D-18-00005**	Dispatch : **17-2-2018**

Galois E (1831). Mémoire sur les conditions de résolubilité des équations par radicaux

Gillies D (1995) (ed.). Revolutions in mathematics. Clarendon Press, Oxford

Grosholz E (2007) Representation and productive ambiguity in mathematics and science. Oxford University Press, New York

Grosholz E, Breger H (eds) (2000) The growth of mathematical knowledge. Springer, Dordercht

Ippoliti E (2006) Demonstrative and non-demonstrative reasoning by analogy. In: Cellucci C, Pecere P (eds) Demonstrative and non-demonstrative reasoning in mathematics and natural science. Cassino University Press, Cassino, pp 307–338

Ippoliti E (2008) Inferenze ampliative. Lulu, Morrisville

Ippoliti E (2016) Ways of advancing knowledge. A lesson from knot theory and topology. In: Sterpetti F, Ippoliti E, Nickles T (eds) Models and Inferences In Science. Springer, Berlin, pp 141–172

Ippoliti E (2018) Heuristic logic. A kernel. In: Danks D, Ippoliti E (eds) Building theories. Springer, Berlin, pp 191–212

Ippoliti, E-Cellucci, C (2016) Logica. Egea, Milano

Kleiner I (1986) The evolution of group theory: a brief survey. Math Mag 59(4):195–215

Kleiner I (2007) A history of abstract algebra. Birkhauser, Boston

Lagrange J (1770). Réflexions sur la résolution algébrique des équations. Nouveaux Mémoires de l'Académie royale des Sciences et Belles-Lettres de Berlin, 134–215

Lagrange J (1808). Traite de la resolution des equations numeriques de tous les degres, 1798, revised in 1808, In: Œuvres de Lagrange, Serret JA (ed), Gauthier–Villars, Paris vol. 8, pp. 1867–1892

Lakatos I (1976) Proofs and refutations: the logic of mathematical discovery. Cambridge University Press, Cambridge

Pengelley DJ (2005). Arthur Cayley and the first paper on group theory. In: Shell-Gellasch A, Jardine D (eds). From Calculus to Computers. MAA, Washington, DC, pp 3–8

Ronan M (2006) Symmetry and the monster. Oxford University Press, New York

Thomas R (2011) Assimilation: not only indiscernibles are identified. In: Cellucci C, Grozholz E, Ippoliti E (eds) Logic and knowledge. Cambridge Scholars Publishing, Newcastle Upon Tyne, pp 363–380

Wussing H (1984) The genesis of the abstract group concept. Dover, New York

Journal : Large 11245	Article No : 9549	Pages : 9	MS Code : TOPO-D-18-00005	Dispatch : 17-2-2018