

THE INTUITIONS OF HIGHER DIMENSIONAL ALGEBRA FOR THE STUDY OF STRUCTURED SPACE*

Ronald BROWN and Timothy PORTER

RÉSUMÉ : Les algèbres de dimensions supérieures libèrent les mathématiques de la restriction d'une notation purement linéaire. Elles aident ainsi à la modélisation de la géométrie et procurent une meilleure compréhension et plus de possibilités pour les calculs. Elles nous donnent de nouveaux outils pour l'étude de problèmes non-commutatifs, de dimension supérieure qui assurent le passage du local au global, en utilisant la notion d'« inverse algébrique de subdivision ». Nous allons exposer comment ces idées sont venues aux auteurs en prolongeant initialement la notion classique de groupe abstrait à celle de groupoïde abstrait, dont la composition n'est que partiellement définie, et qui ajoute une composante spatiale à la théorie habituelle des groupes. La théorie des nœuds nous fournit un exemple en indiquant comment une telle algèbre peut être utilisée pour décrire la structure d'un espace. Le prolongement à la dimension 2 utilise des compositions de carrés dans deux directions et la richesse de l'algèbre qui en résulte est montrée par certains calculs de dimension 2. La difficulté de la transition de la dimension 1 à la dimension 2 est également illustrée par la comparaison de la notion de carré commutatif à celle de cube commutatif – le traitement de cette dernière nécessitant de nouvelles notions. L'importance de la théorie des catégories est expliquée, de même que les possibilités de l'application d'algèbres de dimensions supérieures.

MOTS-CLÉS : algèbres de dimensions supérieures, théorie des nœuds, groupes, groupoïdes, théorie des catégories, connections, méthodes cubiques.

ABSTRACT : Higher dimensional algebra frees mathematics from the restriction to a purely linear notation, in order to improve the modelling of geometry and so obtain more understanding and more modes of computation. It gives new tools for non-commutative, higher dimensional, local to global problems, through the notion of « algebraic inverse to subdivision ». We explain the way these ideas arose for the writers, in extending first the classical notion of abstract group to abstract groupoid, in which composition is only partially defined, as in composing journeys, and which brings a spatial component to the usual group theory. An example from knot theory is used to explain how such algebra can be used to describe some structure of a space. The extension to dimension 2 uses compositions of squares in two directions, and the richness of the resulting algebra is shown by some 2-dimensional calculations. The difficulty of the jump from dimension 1 to dimension 2 is also illustrated by the comparison of the commutative square with the commutative cube – discussion of the latter requires new ideas. The importance of category theory is explained, and a range of current and potential applications of higher dimensional algebra indicated.

KEYWORDS : higher dimensional algebra, knot theory, groups, groupoids, category theory, connections, cubical methods.

* This is a development of a talk by Ronald Brown at the École normale supérieure on 30 May 2001, in the seminar series « Géométrie et Cognition » of Giuseppe Longo, Jean Petitot, Bernard Teissier.

ZUSAMMENFASSUNG : Die mehrdimensionale Algebra befreit die Mathematik von einer rein linearen Notation. Sie ermöglicht eine Modellierung der Geometrie, erleichtert das Verständnis von Rechnungen und stellt dafür eine größere Zahl von Hilfsmitteln zur Verfügung. Sie liefert uns auch neue Werkzeuge für die Untersuchung « lokal-globaler » nicht-kommutativer Probleme, indem der Begriff einer algebraischen Umkehrung der Subdivision eingeführt wird. Wir werden zeigen, wie diese Vorstellungen entstanden sind. Der klassische, abstrakte Gruppenbegriff wurde zu einem abstrakten « Gruppoid » weiterentwickelt, dessen Zusammensetzung nur teilweise definiert ist und bei dem der herkömmlichen Gruppentheorie eine räumliche Komponente hinzugefügt wird. Die Knotentheorie liefert dafür ein Beispiel, denn sie zeigt den Nutzen einer solchen Algebra für die Beschreibung einer Raumstruktur. Bei der Ausweitung auf die zweite Dimension werden in zwei Richtungen Zusammensetzungen von Quadraten benutzt, und der Reichtum der daraus hervorgehenden Algebra wird durch einige Rechnungen in der zweiten Dimension erläutert. Die Schwierigkeit des Übergangs von der ersten zur zweiten Dimension wird durch den Vergleich des Begriffs « kommutatives Quadrat » mit dem Begriff « kommutativer Würfel » erläutert. Um diesen zu behandeln, sind neue Ideen erforderlich. Ferner wird die Bedeutung der Theorie der Kategorien erklärt, und es wird eine Reihe von bereits existierenden und potentiellen Anwendungen der mehrdimensionalen Algebra angegeben.

STICHWÖRTER : mehrdimensionale Algebra, Knotentheorie, Gruppe, Gruppoide, Theorie der Kategorien, Verknüpfungen, kubische Methoden.

RIASSUNTO : L'algebra di dimensioni superiori libera la matematica dai limiti delle notazioni puramente lineari, al fine di migliorare la modellizzazione della geometria ed ottenere una migliore comprensione e migliori strumenti di calcolo. Fornisce inoltre nuovi strumenti per strutture non-commutative, dimensioni superiori e passaggi dal locale al globale, grazie alla nozione di « inverso algebrico della suddivisione ». Spieghiamo il modo in cui gli autori sono pervenuti a queste idee, estendendo in primo luogo la nozione classica di gruppo astratto a quella di gruppoide astratto, nel quale la composizione è definita solo parzialmente, come nei « percorsi di composizione » in cui si aggiunge una componente spaziale alla usuale teoria dei gruppi. Viene usato un esempio della teoria dei nodi al fine di spiegare come una tale algebra possa essere usata al fine di descrivere alcune strutture spaziali. L'estensione a due dimensioni usa la composizione dei quadrati in due direzioni e la ricchezza dell'algebra corrispondente è dimostrata da alcuni conti in due dimensioni. La difficoltà del passare da una a due dimensioni è anche illustrata dal raffronto del quadrato commutativo con il cubo commutativo, la cui discussione richiede concetti nuovi. Si spiega infine l'importanza della teoria delle categorie e lo spazio delle applicazioni attuali e possibili dell'algebra di dimensioni superiori.

PAROLE CHIAVE : algebra di dimensioni superiori, teoria dei nodi, gruppi, gruppoidi, teoria delle categorie, connessione, metodi cubici.

Ronald BROWN, born in 1935, is emeritus professor, « Leverhulme Emeritus Fellow » in the Department of Mathematics, at the University of Wales, Bangor. His research concerns topology, the theory of categories and their application.

Address : University of Wales, Bangor, Department of Mathematics, Dean Street, Bangor, Gwynedd, LL57 1 UT, U.K.

E-mail : r.brown@bangor.ac.uk

Timothy PORTER, born in 1947, is professor of mathematics in the Department of Mathematics, University of Wales Bangor. His research concerns topology, logic, the theory of categories and their application.

Address : University of Wales, Bangor, Department of Mathematics, Dean Street, Bangor, Gwynedd, LL57 1 UT, U.K.

E-mail : t.porter@bangor.ac.uk

We would like to thank Giuseppe Longo for the invitation to give the talk by the first author of which this article is a development, and for suggesting the thrust in terms of explanation of the words in the title. For example he asked a key question : how can an internal notion of algebraic dimension, besides vectorial independence, structure space ?

Since this article is based on a talk in a series including the area of « Cognition », we hope that these attempts to show a rôle of intuition in mathematics will also be useful. Intuition clearly has a cognitive rôle ; it also has an emotional rôle, in displacing fear. It also has a crucial rôle in planning research, and in this area of higher dimensional algebra which Ronald Brown started in 1965 the clarity of some of the intuitions was the force which kept the project going despite a very slow start, and also despite a fairly wide skepticism. However the idea would not go away.

These aspects of intuition, cognition, emotion, and planning are also relevant to the study of the nature of the mathematical process ; one must presume that a « complete » answer to this study would involve so many answers in cognitive science itself that it is clear we can expect only partial answers. It may also be that an adequate cognitive science description of mathematics will in fact need some new mathematics ! The reason for this is that mathematics is also a descriptive language, which may not be translatable.

For example, a photon or electron can be adequately described only in mathematical terms. If instead you try to use ordinary language, then you seem to have paradoxes, as you try to force analogies with events on a different scale. So in the usual two slits experiments, we tend to say in ordinary language that « the photon has to pass through one or the other slits », but this is inadequate to describe what actually happens when an interference pattern builds up in the experiment. It is even more difficult to describe in ordinary language the phenomenon of entanglement, where the states of two particles are linked over a macroscopic distance.

The development of a new mathematics for the description of the cognitive bases of mathematics would correspond to the way in which attempts to give a logical foundation for mathematics led to many new developments in mathematics, and also in logic, so that logic came to be not a foundation for but a part of mathematics (see for example the papers by Longo¹). It would also reflect the capacity for self reference which is an important part of logic. Part of the interest in this seminar for us was that in preparing for giving it and then writing it up, it would develop wider scientific contacts and might spark new

1. LONGO, 1998 and 2001.

ideas on what that kind of new mathematics might be. One would like mathematics to be as effective in, say, neurological sciences as the article by Eugene Paul Wigner² suggests it has been in other sciences. That success has required, as Wigner explains, the development and skillful exploitation of appropriate mathematical concepts.

Mathematics is often thought of in the public mind as concerned with technique and performance, or with problem solving, rather than ideas, and it is perhaps for this reason that the association of mathematics with fear is common. It would be better to see mathematics not as a subject capable of a finished description and account, but as a process, involving refinement of arguments and concepts, and where new fundamental ideas are still possible, even if subject to the usual difficulties of any revolution in science. These new ideas may in fact bypass the apparent and accepted priorities for solving already formulated problems. The development of mathematics is also a sociological process. The book of George Lakoff and Rafael Núñez, *Where mathematics come from. How the embodied mind brings mathematics into being*³, gives further discussion of related issues.

For Ronald Brown, the initial impetus towards higher dimensional algebra came while writing a text on topology in the 1960s⁴. A basic result in topology is known as the Van Kampen theorem : it allows the calculation, for a space which is built up of smaller parts, of an invariant called the « fundamental group ». It seemed useful to generalise this Van Kampen theorem to allow wider calculations, and it turned out as shown in 1967 – in « Groupoids and Van Kampen's theorem⁵ » – that this could be done by generalising the theorem from groups to what in 1926 were called « groupoids ». This notion will be explained later. The notion of abstract group is widely regarded as a fundamental concept in mathematics and physics, as groups are considered the mode for encapsulating notions of symmetry. So the above work raised the question as to the wider advantages of replacing groups by groupoids – a survey on this published in 1987 is « From groups to groupoids. A brief survey⁶ » – which also explains how one stimulus for the notion of groupoids came from extending work of Gauss in arithmetic. Further thought also suggested that the proof of the groupoid Van Kampen theorem could be extended to all higher dimensions if certain key constructions could be carried out. This conjectural theory was mentioned in the introduction to « Groupoids and Van Kampen's theorem⁷ ». There is a general idea here : find algebraic

2. WIGNER, 1991.

3. LAKOFF and NÚÑEZ, 2000.

4. BROWN, 1968 and also 1988.

5. BROWN, 1967.

6. BROWN, 1987.

7. BROWN, 1967.

structures which enable one to describe the behaviour of at least some complex systems in terms of the behaviour of their individual parts.

At the time, this notion of generalising the Van Kampen theorem was an idea of a proof in search of a theorem. All that Brown then knew existed in terms of proper mathematics, with the usual repertoire of definitions, examples, lemmas, theorems, proofs, counterexamples, was a definition and one major example namely the double category of commutative squares in a category as given by Charles Ehresmann in *Catégories et structure*⁸. So it was like having a theme and climax for a play but without the major characters being clear. With further clues being noticed and through fortunate collaborations (Christopher Spencer, Philip Higgins and research students Keith Dakin and Nick Ashley), the characters gradually appeared. A story valid in all dimensions was published in 1981⁹. It was very gratifying to find fifteen years or so from the start that the intuitions had not only sustained the development of a substantial new theory but were also neatly encapsulated in it. The subject is still developing fast. In recent years much wider possibilities for these ideas have become apparent, with applications not only in mathematics but also in physics and computer science, as a web search on « higher dimensional algebra » will show¹⁰.

One of the questions that arose from this success was the following. We normally write mathematics on a line, made up of formulae in which each symbol is related to those to the left and those to the right. (There are some exceptions to this which we will ignore for the moment.) The question is whether this « on a line » is a necessity or whether there are wider possibilities with useful applications.

Here is a quote from John Baez, whose web site and papers¹¹ have a lot on this area :

« [...] not only does higher-dimensional algebra seek to burst free of certain habits of “linear thinking” that tend to go along with symbol string manipulation, it also has been a bit outside the mainstream of mathematics until recently.

« Now, when I speak of “linear thinking” I am not indulging in some vague new-agey complaint against rationality. I mean something very precise : the tendency to focus one’s energy on operations that are easily modelled by the juxtaposition of symbols in a line. The primordial example is addition : we have a bunch of sticks in a row :

|||||

8. EHRESMANN, 1965.

9. BROWN and HIGGINS, 1981a and 1981b.

10. BROWN, 1996 ; STREET, 1996.

11. BAEZ, 2001 and 2003*.

and we say there are “5” sticks and write :

$$1 + 1 + 1 + 1 + 1 = 5.$$

« Fine. But when we have a 2-dimensional array of sticks :

$$\begin{array}{c} ||| \\ ||| \end{array}$$

we are in a hurry to bring the situation to linear form by making up a new operation, “multiplication”, and saying we have 2×4 sticks. This isn’t so bad for plenty of purposes ; it’s not as if I’m against times tables ! But certain things, particularly in topology, can get obscured by neglecting operations that correspond most naturally to higher-dimensional forms of juxtaposition, and Brown’s paper¹² explains some of these, and how to deal with these problems. The point is not to avoid linear notation ; it’s to avoid falling into certain mental traps it can lead you into if you’re not careful¹³ ! »

This questioning of standard usage relates to a comment of Einstein :

« For when I turn to science not for some superficial reason [...] then the following questions must burningly interest me as a disciple of science : What goal will be reached by the science to which I am dedicating myself ? To what extent are its general results “true” ? *What is essential and what is based only on the accidents of development*¹⁴ ? [...] It is therefore not just an idle game to exercise our ability to analyse familiar concepts, and to demonstrate the conditions on which their justification and usefulness depend, and the way in which these developed, little by little [...]¹⁵. »

Recent developments in computer science suggest this « higher dimensional algebra » relates to a « higher dimensional logic » which may have important and practical applications¹⁶.

The great physicist, Paul Dirac, in one of his last addresses, explained his own credo :

« One should allow oneself to be led in the direction which the mathematics suggests [...] [o]ne must follow up [a] mathematical idea and see what its consequences are, even though one gets led to a domain which is completely foreign to what one started with... Mathematics can lead us in a direction we would not take if we only followed up physical ideas by themselves¹⁷. »

12. BROWN, 1992.

13. BAEZ, 1995.

14. Our emphasis.

15. Attributed to Einstein, see *Mathematical Intelligencer*, ser. 12, 2, p. 31.

16. See for example MONTANARI, 2003*, and the references these papers contain.

17. Quoted in FERRIS and FADIMAN, 1991, p. 63.

SPACE

In this section we develop some general ideas on the mathematical notion of space.

In the paper « The space of mathematics. Philosophical, epistemological and historical explorations¹⁸ » of Friedrich William Lawvere, the point is made that a space is both a domain for intensively and extensively variable quantities, as well as an arena for « becoming », where becoming refers to the objective processes of motion, development, and change. By analogy, subjective objects (such as data types) and subjective processes (such as encoding) are sometimes objectified as taking place in space-like categories. In other words, he suggests the important idea that the mathematical notion of space is a repository for the encoding of motion. By « motion » we include « change of data ».

From an evolutionary point of view the ability of an organism to deal with changes in the environment is crucial, so that some method of internal coding or mapping of such changes would develop, and clearly has developed. The great geneticist Theodosius Dobzhansky has written in 1973 a famous paper : « Nothing in biology makes sense except in the light of evolution¹⁹. »

We have moved from the greek view that there was only one geometry, to a much more eclectic approach and in mathematics the word « space » is almost overused. It comes with adjectives such as topological, vector, Hilbert, Euclidean, hyperbolic, and many others. Euclidean space is concerned with the data arising from angle and distance, governed by Pythagoras' theorem ; vector spaces are concerned with directions ; topological spaces are concerned with continuity ; and so on.

Another illustrative idea is in the notion of « phase space » in physics. For example, the motion of a simple pendulum can be represented by a point on a circle, given by the angle of the pendulum to the vertical. But the « phase space » for a double pendulum, which is a short rod fixed to the end of another rod, is described by two circular coordinates and so is what mathematicians call a « torus » and is more familiar as a doughnut or inner tube.

Notice once again that the surface of the torus looks locally flat, just like the Earth, but that globally it is not, and also the torus differs globally from the Earth. One of the aims of mathematics, and in particular the branch called topology, is to show clearly and in a way which leads to calculation in what these differences consist.

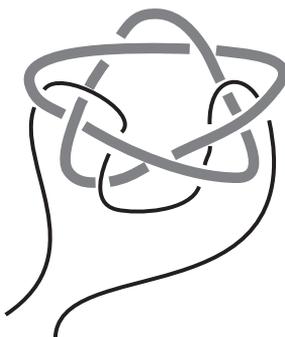
18. LAWVERE, 1992.

19. DOBZHANSKY, 1973.

An important use of the notion of phase space was in celestial mechanics, the study of the motions of the planets and stars. A problem Newton solved was the two body problem : the motion of two bodies under an inverse square law of gravitation. The problem of the motion of three bodies became an important goal in the XIXth century. It gradually became clear that an explicit solution similar to the two body case was not possible, so Poincaré initiated a qualitative study. Each of the three bodies has three coordinates of position and three of velocity, so that makes a total of 18 dimensions. However the motion from a given starting impetus is governed also by the energy equation $E = \text{constant}$, and this restricts the motion to a space M_E of 17 dimensions. A particular motion can be seen as a path in this space M_E . Each point in M_E represents a particular configuration of the three bodies (a total of 9 coordinates) but also has 9 other coordinates that represent the tangents to their curves of motion, thus encoding their change of position. As a point moves smoothly along a path in M_E , the positions and velocities of the bodies in the corresponding configurations change smoothly and their change will be consistent with the energy constraint. So the emphasis was changed from a study of one motion to a study of the phase space M_E which encapsulated all possible motions with a given energy E .

The space M_E has a kind of structure called in mathematics a « 17 dimensional manifold ». An example of a 2-dimensional manifold is a sphere. It is usual to think of the surface of the Earth as a sphere and at a coarse-grained observational level this is a good model. We can use this model in two ways. (a) We can study abstract spheres in order to understand the geometry of the surface of the Earth. (b) We can also use analogy, intuition and our experience of the surface of this planet, to develop and then to study properties of abstract spheres, other 2-dimensional manifolds and, with care, higher dimensional manifolds such as M_E . For instance, it is standard that locally the Earth seems flat, and so we can use flat charts to describe with reasonable accuracy our immediate environment. However the whole Earth cannot be accurately modelled by one flat map, because the Earth is round ! Instead one needs a collection of charts called an atlas and as one travels the Earth one has to switch from one chart to another. The general abstract situation for this is the mathematics of « manifolds ». The set of charts give a particular way of « structuring » our knowledge of the Earth, and one which is of course important for, and indeed derived from, travel, in accordance with Lawvere's dictum.

Here is another illustration of the idea of space from knot theory. Figure 1 includes an illustration of a knot called a pentoil.

FIG. 1. – *Flying around a knot*

In order to study the knot, it is convenient to study the space around the knot, which we call the « knot space » for the pentoil. This is an important point in the development of tools for understanding space. If one is exploring a city or an individual building in a city, it is useful, if not essential, to travel around it. The use of terms such as « maze » or « labyrinth » to describe the structure of the street plan of an old quarter of a city is quite common and when you have traced a tortuous route through narrow alleys arriving once again in front of some distinctive key monument, the sense of achievement, and sometimes relief, is often striking. You have managed to understand something more about the geometry of the city by tracing paths through it. Likewise to understand the complexity of a knot, it is useful to travel (fly ?) in the space around the knot. (Travelling around in the knot itself tells one very little. It is a bit like studying the street plan of the city but never venturing outside of the building you are in !) This space is a bit like a 3-dimensional labyrinth, yet we note that this space looks locally like our ordinary 3-dimensional space. Globally, however, it is quite different.

One of the major problems is how to analyse this difference in ways which lead to deduction and calculation and to methods suitable for a wide range of examples. As with many spatial problems, one possible way is to algebraicise the problem in some way. Algebra, like other parts of the mathematical tool kit, allows for a high degree of certainty in the validity of deductions and calculations. If problems arise, it is usually at the level of the fitness of the algebra to model the feature under study. Once modelled, the logical progression of the algebraic machinery will, hopefully, lead to some greater understanding of the feature. Such has been the experience of many developers or applicers of mathematics and is the reason why mathematics is at the centre of high technology, from engineering and computer science through to biology and cognitive science.

ALGEBRA

In the case of the knot space, we regard flying around the space as probing the space by means of paths, rather like knowledge of the Earth was found by travelers. An important fact is that paths (journeys) can be composed : a path x can be followed by a path y to give a path xy , called the « composition » of the two paths. This is one way in which algebra comes into the study of structured space.

A problem with considering all paths in our knot space is that there are too many of them ! So we try to reduce the complication by classifying the paths, and considering only the classes of equivalent paths. This is analogous to considering species of animals instead of individuals. For our paths, we consider two paths as « the same » if they start and finish at the same points, say p and q , and also if one path can be deformed into the other keeping the end points fixed at p, q . We cannot here go into the precise definition of « deformation » but the essential idea should be clear. We think of the paths as pieces of string with a certain elasticity, so that in the deformation we can stretch or shrink the string to a certain extent. (This analogy with string is not quite right since in a deformation of a path we also allow the path to cross itself.)

The path classes have the remarkable property of forming under composition what mathematicians call a groupoid. A groupoid is a set X of « arrows », a set P of « objects » or « vertices », and two functions $s, t : X \rightarrow P$ called the source and target maps ; we write $x : s(x) \rightarrow t(x)$ for any x in X , and write $X(p, q)$ for elements of X with source p and target q . Further there is a partial composition xy for elements x, y of X defined if and only if the target $t(x)$ of x is the same as the source $s(y)$ of y . This corresponds to the idea that journeys do not have sudden jumps (« quantum leaps ») in the middle. We make further assumptions that the composition is associative (*i.e.* $x(yz) = (xy)z$ whenever both sides are defined) ; that for each p in P there is an element $1_p : p \rightarrow p$ of X which acts as an identity, so that $1_{s(x)}x = x, x1_{t(x)} = x$ for all $x \in X$; and finally for each $x \in X$ there is an inverse $x^{-1} : t(x) \rightarrow s(x)$ such that $xx^{-1} = 1_{s(x)}, x^{-1}x = 1_{t(x)}$. Notice that we use a notation for the operation of standing still. This may seem pedantic but is in fact crucial. This kind of pedantry, the paying attention to simple aspects, is one of the ways in which mathematics progresses. Mathematics was held back for centuries for lack of the number zero, which simply counts the number of things in an empty box.

For those familiar with group theory, these rules can be thought of as axioms for a « group with many identities » and the reason we get round the basic proposition that a group has only one identity is that in a groupoid the composition xy is not always defined, and in fact is defined precisely under the geometric condition $t(x) = s(y)$. This is a key idea for higher dimensional developments, which involve more complex shapes than arrows, and so it is more complicated to say when a composition is defined.

However groupoids were not defined until 1926, and so Poincaré's theory of paths was phrased in terms of groups, that is by looking in a space at the paths that start and end at some single fixed « base point ». This leads to a « fundamental group » for a pointed space, that is for a space with a choice of basepoint. Following this, topologists of the early XXth century sought to express higher dimensional phenomena in terms of generalisations of the fundamental group. In 1932 this direct approach was found to fail since « higher dimensional groups » are less complicated than groups. In fact the mathematician Eduard Čech submitted a paper on « higher homotopy groups²⁰ » to the 1932 International Congress of Mathematicians at Zurich : however two famous topologists Pavel Sergeevich Alexandroff and Heinz Hopf quickly proved that these structures were commutative, that is $xy = yx$ for all x, y , and so can not adequately capture higher dimensional phenomena. We explain this result later. On this ground Čech was persuaded to withdraw his paper, and the importance of his ideas was not realised until three years later.

A construction from geometric data of higher dimensional homotopy groupoids, rather than groups, was achieved by Ronald Brown and Philip John Higgins in 1974, and published in 1978²¹. This gave an impetus to further work with Philip John Higgins and with Jean-Louis Loday²².

The replacement of groups by groupoids has an effect of allowing the set P of objects or vertices of the groupoid to give the notion of a group an additional « spatial component » – this allows for more powerful geometric applications. A group is a groupoid with only one vertex. Studying only groups corresponds say to considering rail journeys from Bangor back to Bangor. But some of us might like to stop off at intermediate stations and the description of the journeys then requires groupoids rather than groups. This can be expressed more mathematically as « subdivision leads from groups to groupoids ». It is partly for this reason that groupoids have powerful applications to local to global problems.

The importance of the passage from groups to groupoids is also suggested by Alain Connes, in *Non-commutative Geometry*²³, where he shows that Werner Heisenberg, while not knowing the formal concept of groupoid, still discovered quantum mechanics by looking at a groupoid of transitions of the hydrogen spectrum rather than a group of symmetry. This confirms the intuitive basis of the concept.

The way we can calculate with groupoids is well seen by considering our classes of paths around a knot. We draw a diagram of the knot, label the overpasses and also give a direction to the knot so that we can say which way round

20. ČECH, 1932.

21. BROWN and HIGGINS, 1978.

22. BROWN and HIGGINS, 1981a and 1981b ; BROWN and LODAY, 1987.

23. CONNES, 1994.

a path goes if it is near to an overpass. We then label a loop class around an overpass by the label of the overpass. But at each crossing we get a relation shown in figure 2 which says that going under the

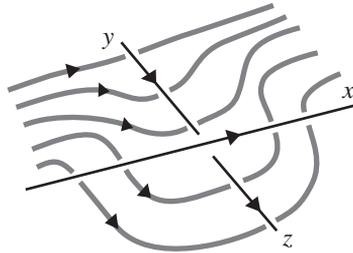


FIG. 2. – Relation at a crossing

overpass y from left to right is equivalent to going under x from top to bottom, then under z from left to right, and finally back under x again from bottom to top : this is much more easily expressed in symbols as :

$$y = x^{-1}zx$$

Thus for the pentoil we get 5 relations, one for each crossing, and we can use them to eliminate all except x, y to get the relation :

$$xyxyxy^{-1}x^{-1}y^{-1}x^{-1}y^{-1} = 1 \tag{1}$$

The left hand side of this equation corresponds to wrapping string around the knot as in figure 3.

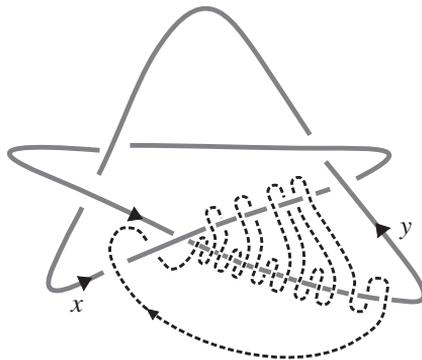


FIG. 3. – Wrapping string around the pentoil

The fact that the left hand side of the equation (1) is « = 1 » means that the string can be undone from the knot without cutting, as was demonstrated in the seminar using a pentoil made of copper tubing. No other formula for tying string on the knot will allow the string to come off the knot. In this way we show how algebra can give to a spatial problem an exact and surprising application. In effect, algebra can show some of the structure of a space.

This knot trick also suggests a typical higher dimensional problem. The equation (1) relates to the fact that the loop of string comes off the knot. The next question is to classify the ways in which it will come off. Is there « essentially » only one way, or are there more than one ? For questions of this type, a higher dimensional approach has been found to give new answers and new computations²⁴.

FROM GEOMETRY TO COMPUTATION

The previous section has given some of the intuitive ideas behind the path or Poincaré groupoid $\pi_1(P)$ of a space P . This groupoid gives rise to path groups $\pi_1(P, p)$ for each point p of P , made up of classes of loops at p , *i.e.* paths from p to itself. The Poincaré group was for a long time considered as more significant mathematically than the Poincaré groupoid. However when Brown was writing a textbook on topology in the mid 1960s it became clear to him that the whole theory made more sense and, possibly as a consequence, the methods of calculation that ensued were more powerful, if the exposition was given in terms of groupoids rather than groups²⁵. The basic reason was that groupoids are better than groups for expressing subdivision. This led him to consider the possible use of groupoids in higher dimensional topology. The excitement of this possibility was that groups play a very significant rôle in mathematics, but it now appeared that they could be just level 1 of a range of new structures for studying higher dimensional phenomena. Thus a naive viewpoint is that n -dimensional geometry requires for its proper expression n -dimensional algebra (as suggested in 1987²⁶). This is part of a general progress in mathematics :

geometry \rightarrow underlying processes \rightarrow algebra \rightarrow algorithms \rightarrow computer implementation

with the aim of enhancing geometrical insight by extensive calculation. Each of these stages requires a different kind of work, and appeals perhaps to different

24. BROWN, 1996.

25. BROWN, 1967.

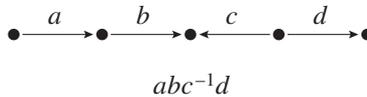
26. BROWN, 1987.

people. The first stage is one that often requires a new range of concepts to describe the « underlying processes », while the algebra can be considered as giving the laws which these concepts have to obey. The advantage of having new algorithms is that immediately you can get new answers, though sometimes computers are needed to do more than the easiest of examples.

Another way of putting the first stage of this process is that to solve some geometric problems may require a new structured language. For the Greeks, this language was the geometry of Euclid. The most notable recent instance of success of this approach of developing a new language to solve problems is the monumental work of Alexander Grothendieck, which laid necessary foundations for the work of Andrew Wiles on Fermat's last theorem. We have a letter from Grothendieck in which he speaks of « the difficulty of bringing new concepts out of the dark », and this suggests that he also saw as an aim for mathematics the development of language for an area, regardless of its success in a well known problem.

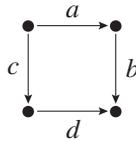
HIGHER DIMENSIONAL ALGEBRA

Now we must get to more details on the intuitions of higher dimensional algebra. Let us start with a basic picture and its algebraic expression :



This is the mathematics of reversible processes : groupoids rather than groups. The formula $abc^{-1}d$ (in which c^{-1} denotes the reverse path) gives a kind of « algebraic inverse » to the subdivision represented by the diagram.

Again, given a square diagram as follows :

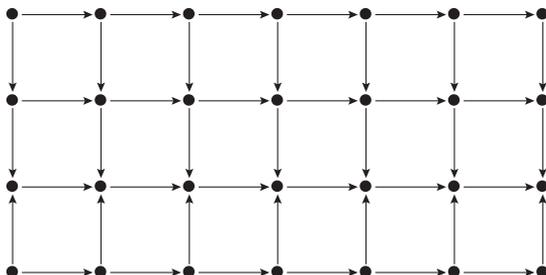


it is easy to say that the square is « commutative » if :

$$ab = cd, \text{ or } a = cdb^{-1}$$

In words, going one way round the square is the same as going round the other way. The formulae given make excellent sense as part of say the theory of groups or of groupoids, and are part of the standard repertoire of mathematics.

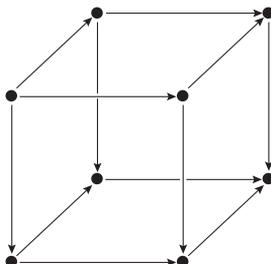
The problem comes when we try to express similar ideas in just one dimension higher. How can one write down algebraically the following picture, where each small square is supposed labelled ?



The diagram shows a square (strictly a rectangle) cut up into little squares but we wish for an algebra which will put them together again. The mathematical word for cutting up is « subdivision », so we can give the slogan of « finding an algebraic inverse to subdivision » in dimension 2. We require an algebra which can express this kind of multiple composition in two directions.

This algebra should also have the property that any multiple composition of commutative squares is commutative. This is a property which is used crucially in the proof of the Van Kampen theorem for the fundamental group or groupoid, that is in the 1-dimensional theorem.

A further problem is to write down algebraically a formula corresponding to one for the above commutative square, but now for the 3-dimensional cube :



What does it mean for the faces of the cube to commute, or for the top face to be the composition in some sense of the other faces ? This question goes to the historical roots of algebraic topology, namely the question of expressing what could be meant by a « composition » of the pieces into which a space may be subdivided. We often use a subdivision to give a structure to a space, cutting it up into small comprehensible pieces, but we wish also to know methods and laws which describe how the space will behave when it is put together.

This is modelled by compositions of cubes and, assuming we know what a commutative cube is, we want to know that any composition of commutative cubes is commutative.

Čech's 1932 work on higher homotopy groups mentioned above can be described as considering a set with two group structures and a compatibility condition called the interchange law. This gave rise to commutative groups, as we can now explain in some detail.

Suppose given a set X with two group structures, say \circ_1, \circ_2 , with identities e_1, e_2 . These structures could be thought of as given by the following pictures :

$$\begin{array}{ccc}
 \begin{array}{|c|} \hline x \\ \hline z \\ \hline \end{array} & & \begin{array}{|c|c|} \hline x & y \\ \hline \end{array} \\
 x \circ_1 z & & x \circ_2 y
 \end{array} \tag{2}$$

These structures should be « compatible » in some way, and the natural condition to impose is that the following diagram of composites

$$\begin{array}{|c|c|} \hline x & y \\ \hline z & w \\ \hline \end{array} \tag{3}$$

can be interpreted in only one way. This is called the « interchange law ».

It is convenient to use a matrix notation. So we write :

$$\begin{bmatrix} x \\ z \end{bmatrix} \text{ for } x \circ_1 z, \quad [x \ y] \text{ for } x \circ_2 y$$

and finally :

$$\begin{bmatrix} x & y \\ z & w \end{bmatrix}$$

for the multiple composition in (3). Notice that these matrices denote not arrays but instead the compositions of the elements formed from the array.

Theorem 4.1 *Let G be a set with two binary operations \circ_1, \circ_2 , each with identities, and satisfying the interchange law. Then the two binary operations coincide, and are commutative and associative²⁷.*

Here is the proof of the theorem.

27. The terms commutative and associative are explained below.

We use some special cases of the interchange law. Let e_1, e_2 denote the identities for the structures o_1, o_2 . We evaluate a matrix composition in two ways yielding easily that :

$$e_1 = \begin{bmatrix} e_1 & e_2 \\ e_2 & e_1 \end{bmatrix} = e_2$$

We write then e for e_1 .

Now we consider the matrix composition :

$$\begin{bmatrix} x & e \\ e & w \end{bmatrix}$$

Interpreting this in two ways yields :

$$x \circ_1 w = x \circ_2 w \tag{4}$$

So we write o for each of o_1, o_2 .

Finally we consider the matrix composition :

$$\begin{bmatrix} e & y \\ z & e \end{bmatrix} \tag{5}$$

and find easily that :

$$y \circ z = z \circ y$$

This is what we mean by saying the binary operation is commutative.

We leave the proof of associativity (that the rule $x \circ (y \circ z) = (x \circ y) \circ z$ always holds) to the reader. This completes the proof.

This seemed from 1932 onwards to be the end of higher dimensional algebra in algebraic topology ! It showed that higher dimensional groups are just commutative groups. This was a tremendous disappointment. Nonetheless the homotopy groups defined by Čech²⁸ were developed by Witold Hurewicz in 1935²⁹ and have become an important area of research in mathematics.

A crucial feature of the above proof is that all the compositions are defined. If you try and write out a similar proof for a set with two groupoid structures you find that all that is proved is that the structure contains a family of commutative groups, but it does not reduce the whole structure to something simpler. This is one of the basic insights for the development of higher dimensional group

28. ČECH, 1932.

29. HUREWICZ, 1935 and 1936.

theory and of higher dimensional algebra. It seems that double groupoids are in some sense « more non-commutative » than groups (or groupoids) and so this does raise the question of their potential future rôle in non-commutative geometry as developed Connes³⁰.

One aspect of the general spirit of higher dimensional algebra is that it is dealing with algebraic structures involving operations defined under geometric conditions. This combination allows a marriage of algebra and geometry. So just as a path is given with two end points, a square is supposed to have four edges and four vertices, and we keep the spirit of diagrams (1) by supposing that the compositions there are defined if and only if the obvious geometric conditions are satisfied. The boundary of a composite square is to be as expected from the geometry : for example, two of the boundary edges of $x \circ_2 y$ are composites of the edges and the other two are edges of x or of y . We also impose the law (2) which is called the interchange law.

Next we need identity squares corresponding to the edges, but in fact we need more as given by the following list.

$$\begin{array}{ccccc}
 \begin{pmatrix} 1 & 1 & 1 \\ & 1 & \end{pmatrix} & \begin{pmatrix} a & 1 & a \\ & 1 & \end{pmatrix} & \begin{pmatrix} 1 & b & 1 \\ & b & \end{pmatrix} & \begin{pmatrix} a & a & 1 \\ & 1 & \end{pmatrix} & \begin{pmatrix} 1 & 1 & a \\ & a & \end{pmatrix} \\
 \square & \dashv & \lrcorner & \lrcorner & \dashv
 \end{array}$$

Here the top line gives the boundaries of the squares and the bottom line gives a notation we shall use. Also any identity edge is written just 1 though in fact there are many identities, corresponding to the points. The first square is a double identity and does not alter a square wherever it is put, provided the composition is defined. The second and third are identities in each direction, so we have equations such as :

$$[x \dashv] = x, \quad \left[\begin{array}{c} x \\ | \\ | \end{array} \right] = x$$

The fourth and fifth are new types of square which are called « connections », and which do alter the boundary of the square. They correspond to the idea that in two dimensions you can not only move in each direction, stop, or turn back, but you can also turn left and right. Thus we begin to see that 2-dimensional algebra holds new possibilities.

So why on earth restrict to dimension 1 ? You will surely know the famous book *Flatland* by Abbott³¹, and where the supposed author visited Lineland,

30. CONNES, 1994.

31. ABBOTT, 1976.

whose inhabitants had very restricted modes of interaction compared to the denizens of Flatland.

When you move up a dimension you get a richer structure which includes the old one and this enables the formulation of new theorems and new proofs. It also turns out that some 1-dimensional phenomena can be best understood from a higher dimensional viewpoint. This is analogous to the fact that the geometry of the Earth is best seen from an orbit outside it.

We now impose the following rules.

Rules 4.2 $[\ulcorner \lrcorner] = \llcorner \llcorner \quad \left[\begin{array}{c} \ulcorner \\ \lrcorner \end{array} \right] = \overline{\llcorner} \quad \left[\begin{array}{cc} \ulcorner & \overline{\llcorner} \\ \llcorner & \ulcorner \end{array} \right] = \ulcorner \quad \left[\begin{array}{cc} \lrcorner & \llcorner \\ \overline{\llcorner} & \lrcorner \end{array} \right] = \lrcorner$

The first two rules are showing the consequences of turning first left and then right, or the other way around, while the other two rules, called the « transport laws », say turning left (right) with arm outstretched is the same as turning left (right).

Using this extra structure of connections we can also answer the question of what a « commutative cube » might mean. One of the problems here is that a cube has six faces and it is therefore not obvious as to how to form an equation of the type « the composition of the even faces is the same as the composition of the odd face ». Using the connections we can write down a sensible formula of the type required ; the cube is to be commutative if its faces satisfy :

$$\begin{array}{|c|c|c|} \hline \overline{\llcorner} & \alpha_3^- & \alpha_2^+ \\ \hline \ulcorner & \alpha_1^+ & \lrcorner \\ \hline \end{array} = \begin{array}{|c|c|c|} \hline \ulcorner & \alpha_1^- & \lrcorner \\ \hline \alpha_2^- & \alpha_3^+ & \overline{\llcorner} \\ \hline \end{array}$$

where α_i^\pm gives the \pm (front and back) faces of the cube α in the direction i . You should check by drawing a 3-cube and labeling the edges that the above formula does give agreement for the edges of the two sides of the above equation. Notice that the cube is 3-dimensional but that the connections are used in an essential way to write down a formula in dimension 2. This tension between different dimensions is one of the features of higher dimensional algebra and is one of its potential strengths, for example in describing and computing with hierarchical systems.

These connections first arose for double groupoids in work of Brown and Spencer in 1976³² and were later developed in all dimensions by Brown

32. BROWN and SPENCER, 1976.

and Higgins in 1979³³. They have been shown crucial in recent work on concurrency in computer science³⁴. It is quite difficult to give analogous formulae for commutative cubes in all dimensions but a method is given in two papers³⁵.

We will now give a flavour of this « two dimensional algebra » by showing a fairly elaborate calculation with these elements, which involves taking a big composition and then reducing it in two different ways.

We would like to prove the equality of the following two 2-dimensional compositions in the case we are given edges a, b, c, d such that the 1-dimensional compositions ab, cd are defined and are equal. This is why the middle rectangles in the following two compositions are well defined.

<table style="border-collapse: collapse; width: 100%; height: 100%;"> <tr> <td style="border: 1px solid black; padding: 5px;"> </td> <td style="border: 1px solid black; padding: 5px;">┌</td> </tr> <tr> <td style="border: 1px solid black; padding: 5px; text-align: center;">c</td> <td style="border: 1px solid black; padding: 5px; text-align: center;">d</td> </tr> <tr> <td style="border: 1px solid black; padding: 5px; text-align: center;"> </td> <td style="border: 1px solid black; padding: 5px;"></td> </tr> <tr> <td style="border: 1px solid black; padding: 5px; text-align: center;">a</td> <td style="border: 1px solid black; padding: 5px; text-align: center;">b</td> </tr> <tr> <td style="border: 1px solid black; padding: 5px;">└</td> <td style="border: 1px solid black; padding: 5px;"> </td> </tr> </table>		┌	c	d			a	b	└		=	<table style="border-collapse: collapse; width: 100%; height: 100%;"> <tr> <td style="border: 1px solid black; padding: 5px; text-align: center;">=</td> <td style="border: 1px solid black; padding: 5px; text-align: center;">a</td> <td style="border: 1px solid black; padding: 5px; text-align: center;">c</td> <td style="border: 1px solid black; padding: 5px;">└</td> </tr> <tr> <td style="border: 1px solid black; padding: 5px;"></td> <td style="border: 1px solid black; padding: 5px; text-align: center;">=</td> <td style="border: 1px solid black; padding: 5px;"></td> <td style="border: 1px solid black; padding: 5px;"></td> </tr> <tr> <td style="border: 1px solid black; padding: 5px;">┌</td> <td style="border: 1px solid black; padding: 5px; text-align: center;">b</td> <td style="border: 1px solid black; padding: 5px; text-align: center;">d</td> <td style="border: 1px solid black; padding: 5px; text-align: center;">=</td> </tr> </table>	=	a	c	└		=			┌	b	d	=	(6)
	┌																								
c	d																								
a	b																								
└																									
=	a	c	└																						
	=																								
┌	b	d	=																						

To prove this we construct a common « subdivision ». One that is appropriate for this case is :

			c		
	□	□	□		┌
	┌	=	a	c	└
		┌	=		
		b	d	=	└
a	└		□	□	□
		b			
					d

33. BROWN and HIGGINS, 1981a and 1981b.
 34. GAUCHER, 2000 ; GOUBAULT, 2000.
 35. AL-AGL, BROWN and STEINER, 2002 ; GAUCHER, 2000.

From this diagram, we may compose parts of the second and third rows using the transport law and then rearrange things once more, getting the left hand side of (6) as indicated below :

$$\begin{array}{|c|c|c|c|} \hline & & c & \\ \hline \square & \square & || & \ulcorner \\ \hline \ulcorner & ab & \bar{=} & \lrcorner \\ a & b & cd & \\ \hline \lrcorner & || & \square & \square \\ \hline & b & & \end{array} \quad d \quad = \quad \begin{array}{|c|c|} \hline & c \\ \hline || & \ulcorner \\ \hline c & d \\ & || \\ a & b \\ \hline \lrcorner & || \\ \hline & b \end{array} \quad d$$

A similar reduction, left as an exercise, gets the right hand side. (This calculation is in the paper by Ronald Brown and Ghafar H. Mosa³⁶, which contains more details on the formal definitions.)

Even if you do not follow the details of this argument, we hope it shows enough to illustrate the richness of higher dimensional algebra and to help build some intuition of how one might more accurately model higher dimensional phenomena by what might be called the algebra of higher dimensional paths.

With this intuition available, let us look briefly at how complex processes such as those found in the biological and cognitive science contexts might be modelled. One possibility might be via some high dimensional manifold, consisting of the various choices of all the observable parameters of the context constrained by various « conservation laws ». Here the geometry will depend strongly on those conservation laws (think back to the three body problem with 18 coordinates constrained by the energy conservation rule). How is one to model conservation of knowledge or belief ? How is one to measure those quantities ? This task seems, to two non-specialists, impossibly difficult, although we will discuss it later. It may however be trying to put the cart before the horse. Our thought experiment would go like this. Imagine such a cognitive phase space could be found as a non-trivial model for some significant aspect of behaviour, knowledge, belief or whatever. How would one study it ? For many spaces initial information on their structure can be gleaned from their « homotopy type », that is, their more combinatorial structure after deformation. It is exactly this sort of structure that the higher dimensional algebra was designed to model, so instead of aiming at a large spatial model as

36. BROWN and MOSA, 1999.

might be appropriate for a model of a purely physical situation, we might look for evidence or ideas for the applicability of models using higher dimensional algebra directly to handle the processes of cognitive science and related areas both of biology and of the computer science/artificial intelligence interface.

At the level of category theory, which is not in itself higher dimensional, there are attempts to model perception categorically in the work of Zippora Arzi-Gonczarowski³⁷ ; considerable discussion of modelling complex hierarchical systems in work of Andree Ehresmann and Jean-Paul Vanbremeersch³⁸ ; and some discussion of issues of categorical pattern recognition by Porter³⁹. These, in a sense, correspond to looking at single processes (analogous to paths) so only use the low dimensional model. If we look at concurrency within computer science, higher dimensional phenomena intervene, Éric Goubault⁴⁰ and Philippe Gaucher⁴¹ model deadlock detection in, for instance, access problems to distributed databases. When modelling true concurrency, the use of Petri nets (due to Carl Adam Petri⁴²) is standard in both the analysis of control systems, for concurrency and for manufacturing systems. Petri nets lead to a special kind of 2-dimensional category and in the papers already referred to⁴³, Fabio Gadducci and Ugo Montanari have shown how going beyond Petri net models, beyond 2-categories, to double categories and their tile models, can allow more complex interacting processes to be studied. (The intuition in their models is very near to the geometric intuition used in the development of our own flavour of higher dimensional algebra even though they were not aware of the constructions coming from the topological context.)

So in a small way, higher dimensional algebra is making its appearance in these areas. Perhaps the use of purely « linear » mathematics (in the sense of « written on a line » or « 1-dimensional ») may soon be augmented in highly significant ways by its (underdeveloped as yet) higher dimensional relatives. If this is so then the above equations raise again the important point of subdivision as an inverse for composition.

If one analyses a number, say, 45, one has little choice as to how to « deconstruct » it multiplicatively : $45 = 3 \times 3 \times 5$. We have argued elsewhere that this extremely simple model can tell us a lot about the difficulties that face reduction to simpler parts⁴⁴. Here our point is slightly different : looking at

37. ARZI-GONCZAROWSKI and LEHMANN, 1998.

38. EHRESMANN and VANBREMEERSCH, 2001.

39. PORTER, 1994 ; CORDIER and PORTER, 1989.

40. GOUBAULT, 2000.

41. GAUCHER, 2000.

42. PETRI, 1973.

43. GADDUCCI and MONTANARI, 2000 ; MONTANARI, 2003*.

44. In a paper, in preparation, by BROWN (Ronald), PATON (Ray) and PORTER (Timothy), « Categorical hierarchical models for cell systems ? ».

ways of decomposing a higher dimensional « multi process » into simpler processes raises questions such as :

(a) Are there unique « atomic multi processes », (analogous to prime numbers) ?

(b) What decomposition rules apply ? (We know $45 = 3 \times 3 \times 5$ but it is also $5 \times 3 \times 3$. Order does not matter. For 2-dimensional decompositions, we know the equation (5) above applies and the analysis of Al-Agl, Brown and Steiner⁴⁵ can be interpreted as giving a complete set of rules for one interpretation of decompositions in higher dimensions. However the tile compositions used in the papers cited in the publications web page of Montanari⁴⁶ are more general.)

(c) If one decomposes into some choice of « generators » *i.e.* some generating « atomic » multi processes, how can one « rewrite » for another choice ?

(The choice of generators corresponds to a choice of language, and a choice of the logic for handling that language corresponds to the decomposition rules. Those decomposition rules can be interpreted as « rewrite » rules in the sense of rewriting theory but, interestingly, rewrite theory leads back into higher dimensional algebra⁴⁷. Is this by chance or is it saying something deep about our cognitive processes when « modelling » ?)

Consider composition of arrows in a category or groupoid. We tend to write this as an equation :

$$x \circ y = z$$

identifying z as the outcome of applying the composition process to the arrows x and y :

$$A \xrightarrow{x} B \xrightarrow{y} C \quad \text{gives} \quad A \xrightarrow{z} C$$

but the use of « equals » conveys the wrong impression since the left hand side is an expression and the right hand side is an element. It seems that the equation may be causing an example of the old error of identifying the end with the means, the result of a process with the process itself. Nearly always it is the process that ends up being the more important aspect so we will replace $x \circ y = z$ by some notation such as :

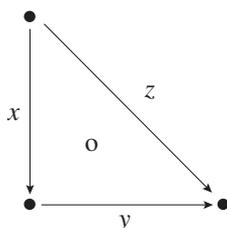
$$x \circ y \Rightarrow z$$

45. AL-AGL, BROWN and STEINER, 2002.

46. MONTANARI, 2003*.

47. See, for instance, BROWN and SPENCER, 1976 ; EHRESMANN and VANBREMEERSCH, 2001.

indicating x followed by y « rewrites to » z . The arrow formalism can even be written in a geometrical fashion :



Here o is the black box (process) which does the calculation.

Now the geometrical views of processes and « processes of processes » yield higher dimensional geometries, and so higher dimensional algebra. Here are two quotes on this, which show a geometricisation of ideas of processes, and seem quickly to capture some key ideas. They are taken with permission from two out of forty-three personal statements for a National Science Foundation research grant proposal by John Baez and Peter May on « n -categories, foundations and applications » :

« The idea at the center of my work since 1991 is that a collection of processes cooperating together to run some program can be modelled as a geometric shape or “space of configurations”. There is a need to characterise the shapes coming from the semantics of concurrent programs . » (Éric Goubault, 1991)

« An “operation” is something which takes several things as input and produces one thing as output ; we naturally picture this as a black box with the inputs as wires coming into the top and the output as a wire coming out at the bottom. Operations can be composed, and then the input is a planar tree of black boxes wired together and giving output from a single box. The process can be iterated leading to shapes of ever higher dimensions. » (Tom Leinster, 2001)

The above has given some indication of one strand in the development of higher dimensional algebra. Other workers have taken up the ideas in different ways and for different purposes – a web search on « higher dimensional algebra » shows applications not only in mathematics but also in physics and computer science. In particular the work of Baez⁴⁸ looks for higher dimensional algebra to give a handle on unified field theories. One aspect of this is that the notion of equality needs to be weakened, since two structures are rarely equal but they are often « isomorphic ». So one gets new kind of algebraic structures

48. BAEZ, 2003*.

with laws which hold only « up to isomorphism ». This theory is still being worked out, and there are many candidates for it⁴⁹.

LANDMARK THEOREMS

What from the point of view of this article are major difficulties and results in the area of higher dimensional algebra ?

In dimension 1 the shapes which could form a basis for a geometric view of algebra consist only of a point and a line segment with an arrowhead imposed. In dimensions more than 1 the number of usable figures is infinite.

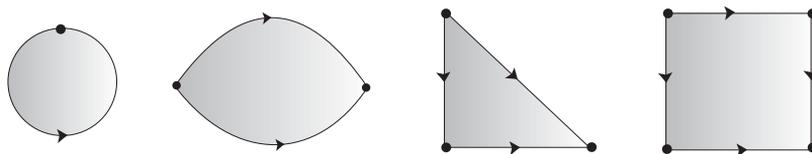


FIG. 4. – *Shapes*

The figure 4 shows the first few in dimension 2 : disc, globe, triangle (or 2-simplex), square. Handling this variety of shapes in an algebraic manner is one of the chief problems of developing the theory, as is showing that in some cases different descriptions are equivalent. The use of globes and simplices is central to much work in this area, as shown in the survey of Ross Street⁵⁰.

A further problem is to develop invariants analogous to the fundamental group or groupoid in these higher dimensions. In fact John Henry Constantine Whitehead in the 1940s developed work based on discs (giving a new concept of crossed module), and the relation of this with work based on squares (*i.e.* double groupoids) was a major feature of the generalisations of the Van Kampen theorem to higher dimensions⁵¹. The development of higher dimensional versions of this theory was a major task⁵². Michael Atiyah⁵³ suggests that key themes of XX^{th} century mathematics have included local to global problems, non commutative methods, and higher dimensions. Higher dimensional algebra combines these three themes, and in particular, the work on higher dimensional

49. See, for example, LEINSTER, 2002.

50. STREET, 1996.

51. BROWN, HARDIE, KAMPS and PORTER, 2002 ; BROWN and HIGGINS, 1978.

52. BROWN and HIGGINS, 1981a and 1981b ; BROWN and LODAY, 1987.

53. ATIYAH, 2001.

Van Kampen theorems has given a lead on the types of algebraic structures suitable for such combined work.

It is important that the cubical work of Brown and Higgins in « The algebra of cubes⁵⁴ » was a stimulus to work in concurrency theory⁵⁵. The start of the use of cubical methods was due to Vaughan Pratt⁵⁶ and is based on the idea that the use of a number say N of processors should be represented by time of N dimensions. So one has to develop higher dimensional analogues of paths, in terms of multiple compositions of small pieces.

There is also a large theory of what are called « weak n -categories » which is still being worked out – see again the account by Leinster⁵⁷. This theory seems to be essential for some applications in physics⁵⁸ and in some areas of computer science and geometry. Thus there is a great « buzz » of developments and prospects.

Also, there has to be a way of constructing a space from the algebra, like constructing a real building from a number of diagrammatic charts. One looks at small pieces of algebra and constructs a space from that, and then shows how larger pieces of algebra lead to larger spaces. There is a lot of active research in this area (which is known as « classifying space theory »).

COMPUTATIONAL ISSUES

A well established model of computation is that of Turing, where a basic tool is a tape on which symbols are printed or erased. In this model, the array given in our introduction has to be translated into a list of lists, *i.e.* we have to replace the array by 8 or by 2×4 . One of the questions is whether there are other possible modes of computation in which the higher dimensional structure is not lost.

There is an important point here related to the confidence we have in mathematics. There are several reasons for this confidence. One is the open nature of mathematics, so that all the results and proofs are available for inspection. In this way sometimes quite long standing errors have been discovered. Another is its abstract nature, so that particular systems such as vector spaces, or homology theories, have evolved over a long period of time, and are tested in many instances. Another feature here is the modular nature of mathematics.

54. BROWN and HIGGINS, 1981a.

55. GAUCHER, 2000.

56. PRATT, 2003*.

57. LEINSTER, 2002.

58. BAEZ, 2001.

A less well recognised feature is its structural nature. Thus mathematicians are often « certain » of a proof even when not all the logical details have been checked – indeed, mathematics is not usually written in a way which enables the details to be checked at what one might call a machine logical level. But would such a check give certainty ? A contrast could be given between describing a route to a railway station by listing a few turns and instead by listing all the cracks, etc., in the pavements on that route. The former method is more certain than the latter – within limitations, such as avoiding manholes ! We do need theorem provers, which could help analogously to finding a route through a maze, but they need to work at the appropriate conceptual level. So the problem can be summarised : how can we reflect in computation the conceptual, modular, hierarchical structure of mathematics ? How can one do structural mathematical computations (if that is the right phrase) ? If the structural features by which we understand the mathematics of an algorithm are lost in the passage to a program, then the programmer also may make errors which are hard to detect. It will be very interesting to see if the development of new type theories and the languages which use them can solve this problem and still work in an efficient way to express complex algorithms to solve real computational problems.

Another point to make here is that there are very well used Computer Algebra Systems which usually give a very nice interface for calling some efficient algorithms. There are three potential problems with these : (a) sometimes to cope with a new problem, or to improve the methods, these algorithms really need to be tweaked or modified, and this can be difficult to accomplish ; (b) the computer algebra system may carry overheads which make it difficult to achieve the greatest efficiency ; (c) if for commercial reasons the mathematics of the algorithms is hidden, what confidence should the user have in the results the algorithms provide ?

A great contribution to the progress of mathematics is exposition and explanation, so that it is not just a question of doing a calculation to verifying something, but also of knowing why something is true. That is, mathematicians are looking also for higher levels of explanation, however one defines such levels, so that a result has to be true because of the way it fits with general structural concepts. This theme is related to the role of intuition in the development of mathematics, and perhaps also to the role of beauty, as has been emphasised by many scientists.

Computers have proved very useful in applying algorithms to complex situations, but in higher dimensional algebra there are not so many of these algorithms. The finding of these for doing calculations that are needed is one of the challenges.

Higher dimensional ideas have proved useful for formulating and proving theorems, but turning these into computational tools has required a translation into a « linear » framework. Thus different views of the « same » kind of object

have been necessary to turn theory into practice. It is hoped that this method will prove more widely useful.

Is the linear model of computation adequate for mathematics ? Would a truly concurrent programming style use a higher dimensional concept of computation ?

There are a number of indications that more is needed. Several papers by Gadducci and Montanari⁵⁹ give a 2-dimensional model using double categories for problems related to processes and to concurrency. This also reacts with models for modal logics, which deal with the problems of conservation of knowledge, and of belief revision. The study of computation within such models and within quite general multiagent systems would seem to need some of the techniques and conceptual models of higher dimensional algebra.

It is the aim of this article to point a possible way and to stimulate further investigation. The basic intuitions are very strong and as Brown wrote provocatively in 1987 « n -dimensional geometry requires n -dimensional algebra⁶⁰ », thus suggesting a broader aspect of the 1982 article « Higher dimensional group theory⁶¹ ».

Ronald BROWN and Timothy PORTER*
(February 2003).

59. See, for example, GADDUCCI and MONTANARI, 2000.

60. BROWN, 1987, p. 128.

61. BROWN, 1982.

* We would like to thank : Giuseppe Longo for the invitation to lecture to his seminar ; Posina Venkata Rayudu of the National Brain Research Centre, Gurgaon, India, for helpful comments ; and Andree Ehresmann for helpful comments and additional references – readers interested in applications of category theory in biology should explore the web site <<http://perso.wanadoo.fr/vbm-ehr/>>. The works of Charles Ehresmann in category theory and of John Henry Constantine Whitehead in topology have been a strong influence in the development of this theory. We are grateful also to Laurie Catteeuw for her help in preparing the paper for this publication.

LIST OF REFERENCES

- ABBOTT (Edwin A.), 1976, *Flatland. A romance of many dimensions*, 1st ed. London, 1884, here Cutchogue, NY, Buccaneer Books. The text is also available at <<http://www.alcyone.com/max/lit/flatland/>>.
- AL-AGL (Fahd), BROWN (Ronald) and STEINER (Richard), 2002, « Multiple categories. The equivalence between a globular and cubical approach », *Advances in Mathematics*, 170, p. 71-118.
- ARZI-GONCZAROWSKI (Zippora) and LEHMANN (Daniel), 1998, « The category of artificial perceptions », *Annals of Mathematics and Artificial Intelligence*, 23, p. 267-298. (The site <<http://users.actcom.co.il/typographics/zippie/>> contains other relevant work by the same authors.)
- ATIYAH (Michael), 2001, « Mathematics in the XXth century. Geometry versus Algebra », *Mathematics today*, ser. 2, 37, p. 47-49.
- BAEZ (John), 1995, « This week's finds in mathematical physics », *Week 53*, 18 May (this text is also available at <<http://math.ucr.edu/home/baez/twf.html>>).
- BAEZ (J.), 2001, « Higher-dimensional algebra and Planck-scale physics », in CALLENDER (Craig) and HUGGETT (Nick), eds., *Physics meets philosophy at the Planck scale*, Cambridge, Cambridge University Press, p. 177-195.
- BAEZ (J.), 2003*, <<http://math.ucr.edu/home/baez/papers.html>> (date of consultation).
- BROWN (Ronald), 1967, « Groupoids and Van Kampen's theorem », *Proceedings of the London Mathematical Society*, ser. 3, 17, p. 385-401.
- BROWN (R.), 1968, *Elements of modern topology*, Maidstone, McGraw Hill.
- BROWN (R.), 1982, « Higher dimensional group theory », in BROWN (Ronald) and THICKSTUN (Tom L.), eds., *Low dimensional topology*, Cambridge, Cambridge University Press (London Mathematical Society Lecture Note Series, 48), p. 215-238.
- BROWN (R.), 1987, « From groups to groupoids. A brief survey », *Bulletin London Mathematical Society*, 19, p. 113-134.
- BROWN (R.), 1988, *Topology. A geometric account of general topology, homotopy types and the fundamental groupoid*, Chichester, Ellis Horwood.
- BROWN (R.), 1992, « Out of line », *Royal Institution Proceedings*, 64, p. 207-243.
- BROWN (R.), 1996, « Higher dimensional group theory », web article available at <<http://www.bangor.ac.uk/~mas010/hdaweb2.html>>
- BROWN (Ronald), HARDIE (Keith), KAMPS (Heiner) and PORTER (Timothy), 2002, « A homotopy double groupoid of a Hausdorff space », *Theory and Application of Categories*, 10, p. 71-93.
- BROWN (Ronald) and HIGGINS (Philip John), 1978, « On the connection between the second relative homotopy groups of some related spaces », *Proceedings of the London Mathematical Society*, ser. 3, 36, p. 193-212.

- BROWN (R.) and HIGGINS (P. J.), 1981a, « The algebra of cubes », *Journal of Pure and Applied Algebra*, 21, p. 233-260.
- BROWN (R.) and HIGGINS (P. J.), 1981b, « Colimits of relative homotopy groups », *Journal of Pure and Applied Algebra*, 22, p. 11-41.
- BROWN (Ronald) and LODAY (Jean-Louis), 1987, « Van Kampen theorems for diagrams of spaces », *Topology*, 26, p. 311-335.
- BROWN (Ronald) and MOSA (Ghafar H.), 1999, « Double categories, 2-categories, thin structures and connections », *Theory and Applications of Categories*, 5, p. 163-175.
- BROWN (Ronald) and SPENCER (Christopher B.), 1976, « Double groupoids and crossed modules », *Cahier de topologie et géométrie différentielle catégorique*, 17, p. 343-362.
- ČECH (Eduard), 1932, « Höherdimensionale Homotopiegruppen », *Verhandlungen des Internationalen Mathematiker-Kongresses Zurich*, Zurich-Leipzig, O. Füssli Verlag, vol. II, p. 203.
- CONNES (Alain), 1994, *Non-commutative geometry*, transl. Sterling K. BERBERIAN, San Diego (CA)-London, Academic Press.
- CORDIER (Jean-Marc) and PORTER (Timothy), 1989, *Shape theory. Categorical methods of approximation*, Chichester, Ellis Horwood (Mathematics and its Applications).
- DOBZHANSKY (Theodosius), 1973, « Nothing in biology makes sense except in the light of evolution », *The American Biology Teacher*, 35, march, p. 125-129.
- EHRESMANN (Charles), 1965, *Catégories et structure*, Paris, Dunod.
- EHRESMANN (Andree) and VANBREMEERSCH (Jean-Paul), 2001, « Emergence processes up to consciousness using the multiplicity principle and quantum physics », *American Institute of Physics Conference Proceedings*, August. (See also other papers on <<http://perso.wanadoo.fr/vbm-ehr/>>.)
- FERRIS (Timothy) and FADIMAN (Clifton), eds., 1991, *The World Treasury of physics, astronomy and mathematics*, Boston (MA)/London, Little/Brown and Co.
- GADDUCCI (Fabio) and MONTANARI (Ugo), 2000, « The tile model », in PLOTKIN (Gordon), STIRLING (Colin) and TOFTE (Mads), eds., *Proof, language and interaction. Essays in honour of Robin Milner*, Cambridge (MA)-London, MIT Press.
- GAUCHER (Philippe), 2000, « Homotopy invariants of higher dimensional categories and concurrency in computer science », *Mathematical Structures in Computer Science*, 10, p. 481-524.
- GOUBAULT (Éric), 2000, « Geometry and concurrency. A user's guide », *Mathematical Structures in Computer Science*, 10, p. 411-425.
- HUREWICZ (Witold), 1935 and 1936, « Beiträge zur Topologie der Deformationen », *Nederlandsche Akademie van Wetenschappen*, ser. A, 38, p. 112-119, 521-528, and 39, p. 117-126, 213-224.
- LAKOFF (George) and NÚÑEZ (Rafael), 2000, *Where mathematics comes from. How the embodied mind brings mathematics into being*, New York, NY, Basic Books.
- LAWVERE (Friedrich William), 1992, « The space of mathematics. Philosophical, epistemological and historical explorations », in *Categories of space and of quantity, International Symposium on structures in mathematical theories*, San Sebastian, Spain, 1990, Berlin, De Gruyter, p. 14-30.
- LEINSTER (Tom), 2002, « A survey of definitions of n -category », *Theory and Applications of Categories*, 10, p. 1-70.

- LONGO (Giuseppe), 1998, « Mathematics and the biological phenomena », in *Proceedings of the International Symposium on Foundations in Mathematics and Biology. Problems, projects and interactions*, invited lecture, Pontifical Lateran University, Vatican City, to appear.
- LONGO (G.), 2001, « The reasonable effectiveness of mathematics and its cognitive roots », in BOI (Luciano), ed., *New interactions of mathematics with natural sciences and the humanities*, Berlin-New York, Springer, to appear.
- MONTANARI (Ugo), 2003*, Publications web page, <<http://www.di.unipi.it/ugo/ugo.html>>.
- PETRI (Carl Adam), 1973, « Concepts of net theory », in *Mathematical Foundations of Computer Science*, Proceedings of Symposium and Summer School, Mathematical Institute of the Slovak Academy of Sciences, 3-8 Sept. 1973, High Tatras, Czechoslovakia, p. 137-146.
- PORTER (Timothy), 1994, « Categorical shape theory as a formal language for pattern recognition », *Annals of Mathematics and Artificial Intelligence*, 10, p. 25-54.
- PRATT (Vaughan), 2003*, Guide to Papers on Chu Spaces, <<http://chu.stanford.edu/guide.html>>.
- STREET (Ross), 1996, « Categorical structures », in HAZEWINDEL (Michiel), ed., *Handbook of algebra*, Amsterdam-Oxford, Elsevier, vol. I, p. 529-577.
- WIGNER (Eugene Paul), 1991, « The unreasonable effectiveness of mathematics in the natural sciences », 1st ed. 1960, *Communications in Pure and Applied Mathematics*, 13, p. 1-14, here repr. in FERRIS and FADIMAN, 1991, p. 526-540.