Gonit Sora (http://gonitsora.com)
Invited Article

COMMUTATIVITY AND NON-COMMUTATIVITY IN MATHEMATICS

Debraj Chakrabarti[1]

Author Information. Dr.Debraj Chakrabarti is an Assistant Professor of Mathematics in Central Michigan University, USA. He is a very good person, a wonderful teacher and a brilliant mathematician. He had done B.Tech from IIT Kharagpur, India and thereafter he completed his masters and PhD in Mathematics from University of Wisconsin, USA. Before joining Central Michigan University, he had served as a faculty in IIT Bombay and also in TIFR-CAM (TIFR Centre for Applicable Mathematics, Bengaluru). More information about him can be found in his homepage (http://people.cst.cmich.edu/chakr2d/). Dr. Chakrabarti has kindly consented to provide Gonit Sora his most valuable time in form of this interesting article.

Keywords: commutativity, quaternions, commutator of an element, partial derivative

Mathematics Subject Classification. 13-01, 20A99

1. Introduction

A patient goes to a dentist with a bad toothache, and the first thing she does is to inject some anaesthetic into the patient's gums. Then she pulls out the wisdom tooth, which was causing the pain. Imagine what would happen if the order of the dentist's actions was reversed: first, the tooth is extracted, and then the anaesthetic is given. The patient's experience would be very different! We students of mathematics would say: the two operations do not commute. The question of commutativity is as important in mathematics as it is in the real world. Given a binary operation $X \times X \to Y$ on a set X with values in Y, the first question to ask about it is whether it is commutative (if further $Y = X$, the next question should be whether the operation is associative.)

The common binary operations in mathematics fall roughly in three classes. First the commutative ones: union and intersection of sets, sum and product of numbers, the sum of vectors, the dot product of vectors, construction of the straight line through two given points on the plane, and so on. The second class consists of operations which are not commutative, but still the role of the two inputs of the binary operation are equivalent, and there are interesting situations when the binary operation fails to commute: multiplication of square matrices, the cross product of vectors, the composition of functions (from a set X to itself) being the typical ones. Finally, operations of the third type are very highly non-commutative: subtraction, division, construction of the circle with centre at one point and passing through another are examples of this third type. In these, the two objects play different roles in the operation, and the operation commutes only in trivial situations or never at all. For a phenomenon which is commutative in nature, we can fruitfully ask: *does it have a non-commutative analogue?* This question often leads to interesting results and investigations. In non-commutative situations of the second type described above, it often makes sense to ask : *what is the degree of non-commutativity of this operation?* Is it possible to quantify in a precise way the deviation from commutativity? In this note, we would like to show how these questions are answered in three concrete

10 March, 2014.

[1] Department of Mathematics, Central Michigan University, Mt. Pleasant, MI 48859, USA

situations. This will lead us to a short tour of some pieces of mathematics where commutativity plays a role, relate some interesting facts about them, have some fun, and leave the interested reader with further questions.

2. The Quaternions

Our first example is a non-commutative analogue of the most important notion of mathematics: *Number.* Through the ages, the notion of number has been successively refined and extended: starting from the positive integers, then to integers (positive and negative), then to rationals, and then to the real numbers, and finally by the nineteenth century, the complex numbers. Algebraically, the complex numbers $\mathbb{C}$ form a *field*, but it is also a real vector space of dimension 2. Denoting the natural basis of $\mathbb{R}^2$ as $1 = (1, 0)$ and $i = (0, 1)$, and tolerating a little abuse of notation (the symbol 1 has two meanings now!), we can represent a complex number z as $z = x \cdot 1 + y \cdot i$. This allows us to visualize the complex number z as the point (x, y) in $\mathbb{R}^2$. The multiplication in $\mathbb{C}$ is completely known as soon as you know that $i^2 = -1$. One can easily check that the multiplication so defined turns the vector space $\mathbb{R}^2$ into a field. One can ask: are there non-commutative generalizations of the complex numbers? We make this vague question precise in the following way. A vector space V over $\mathbb{R}$ is said to be an $\mathbb{R}$-*algebra*, if we are also given a "multiplication" $V \times V \to V$. All we need from the multiplication is that it is linear in each factor – for x, y, x', y' in V and real numbers a, b, a', b' the product $(ax + by)(a'x' + b'y')$ is exactly what it should be: $(aa')(xx') + (ab')(xy') + (ba')(yx') + (bb')(yy')$. Note that in a general $\mathbb{R}$-algebra it is not assumed that the multiplication is either commutative or associative. An $\mathbb{R}$-algebra is called a *division algebra over* $\mathbb{R}$ if further the set $V \setminus \{0\}$ of nonzero vectors in V is a group under the multiplication operation, and the product of any element of V with 0 is 0. We denote the identity element of the group $V \setminus \{0\}$ by 1, again allowing ourselves abuse of notation. Note that elements of V of the form $a.1$ where $a \in \mathbb{R}$ form a copy of $\mathbb{R}$ within V, so that V can be thought of containing $\mathbb{R}$, and therefore a generalization of real numbers. Clearly, $\mathbb{R}$ and $\mathbb{C}$ are division algebras over $\mathbb{R}$. The vector space of 2×2 matrices over $\mathbb{R}$ is an $\mathbb{R}$-algebra under the standard multiplication, though it is not a division algebra, as there are nonzero matrices with no inverse. We can now make precise the question posed in the beginning of this paragraph: are there any other division algebras over $\mathbb{R}$ except our old friends $\mathbb{R}$ and $\mathbb{C}$? This question was answered by the great Irish mathematician Sir William Rowan Hamilton (1805–1865) . His initial motivation was geometric: just as in the complex plane, rotations about the origin are represented by multiplication by complex numbers of unit modulus, is it possible to represent three-dimensional rotations by means of some multiplicative structure defined on $\mathbb{R}^3$? Hamilton tried in vain to construct such a product operation on $\mathbb{R}^3$. Each morning at breakfast, his sons Archibald and William Edward would ask him, "Well Papa, can you multiply triples?", to which he answered, "No, I can only add and subtract them."

Then on 16th October 1843, as he was taking a stroll with his wife on the Brougham Bridge on the Royal Canal in Dublin, Sir William had a moment of inspiration: he needed to modify his question in order to answer it! First, the product would be defined in $\mathbb{R}^4$ rather that $\mathbb{R}^3$, and second : *the product would not commute.* He called his new four-dimensional numbers *quaternions.* The set $\mathbb{H}$ of quaternions is just the vector space $\mathbb{R}^4$ with a product structure, defined in terms of the standard basis elements $1 = (1, 0, 0, 0), i = (0, 1, 0, 0), j = (0, 0, 1, 0)$ and $k = (0, 0, 0, 1)$ by the rules $i^2 = j^2 = k^2 = -1$ and $ij = k, jk = i, ki = j$. Hamilton immediately inscribed these equations on a stone on the Brougham Bridge. These relations suffice to find any other product by associativity, e.g. $ji = (ki)i = k(i^2) = k(-1) = -k$, which already shows that the multiplication on $\mathbb{H}$ is not commutative. One can check that with this product structure, the vector space $\mathbb{H}$ is a four-dimensional division algebra over $\mathbb{R}$. What is more remarkable, $\mathbb{H}$ also contains $\mathbb{C}$ inside it: for example look at quaternions of the form $a \cdot 1 + b \cdot i$, where $a, b \in \mathbb{R}$ (the interested reader is invited to find infinitely many other copies of $\mathbb{C}$ inside $\mathbb{H}$!)

Many properties of the complex numbers generalize to the quaternions. For example, if x, y are quaternions, we have $|xy| = |x|\,|y|$, where $|\cdot|$ denotes the length of a vector in $\mathbb{R}^4$ in the euclidean norm. We write a general quaternion as $x = x_0 + x_1 i + x_2 j + x_3 k$. Note the abuse of notation (similar to $\mathbb{C}$) by which we have written $x_0 \cdot 1$ as x_0, identifying the quaternion $x_0 \cdot 1 + 0 \cdot i + 0 \cdot j + 0 \cdot k$ with the real number x_0. We define $\overline{x} = x_0 - (x_1 i + x_2 j + x_3 k)$ (the *quaternionic conjugate* of x), and then in analogy with the complex numbers we have the relation $x\overline{x} = |x|^2$, and the multiplicative inverse is given by $x^{-1} = |x|^{-2}\,\overline{x}$. It follows that if u

is a fixed quaternion, and we identify the point $x = (x_1, x_2, x_3)$ with the quaternion $x_1 i + x_2 j + x_3 k$, then $x \mapsto uxu^{-1}$ defines a rotation of $\mathbb{R}^3$, and in fact each orientation-preserving rotation of $\mathbb{R}^3$ is of this form, thus solving the original problem of Hamilton of representing three dimensional rotations algebraically. The reader who is not familiar with quaternions is invited to try to prove some of the statements made in this paragraph! When they were introduced, Quaternions represented a revolutionary advance in mathematics. Here was a new system of numbers, in which the product of two things depended on the order of multiplication, yet with properties very similar to those the complex numbers. Quaternions opened the way to the study of abstract algebraic systems for their own sake. But then, *can one go further?* This was answered in 1877 by the a theorem of the German mathematician Ferdinand Georg Frobenius (1849 - 1917): *there are only three division algebras over* $\mathbb{R}$*, which are* $\mathbb{R}$, $\mathbb{C}$ *and* $\mathbb{H}$. On the other hand, one *can* go further, if one is ready to drop the assumption of associativity on the multiplication operation: but this essay is about the commutative property!

Some other important properties of complex numbers do not generalize to the quaternions. For example, the fundamental theorem of algebra (that a polynomial of degree n has exactly n roots) is no longer valid: the polynomial $z^2 + 1$ has infinitely many roots in $\mathbb{H}$ including $\pm i, \pm j$ and $\pm k$. It is an interesting exercise to find all roots! A more serious failure is that the elegant methods and results of complex analysis (theory of analytic functions of a complex variable) have no direct quaternionic analogues.

3. Group Theory

Recall that a commutative group is also called abelian, after the Norwegian mathematician Neils Henrik Abel (1802 –1829.) Many important groups are nonabelian. An interesting example (related to the quaternions!) is the group of rotations of a rigid body in three dimensional space. Choose two straight lines in space intersecting at right angles, and call them x- and y axes. Denote by X the operation of rotating a body through a right angle about the x-axis, and by Y the operation of rotating it through a right angle about the y-axis. It is not difficult to see that $XY \neq YX$. The reader is urged to stop reading this article, close this copy of Craze@Math, and after mentally choosing the x and y-axes passing through the book and intersecting at a point inside it, perform the operations X and then Y to obtain the action of YX. Again going back to the initial position one can perform Y first and X next in order to obtain XY, and check that the end results are different. For a group G, is there a way of understanding how much it deviates from being abelian? Trying to answer this question leads to some important constructions in Group Theory. Given two elements g, h in a group g, denote by $[g, h]$ the element $ghg^{-1}h^{-1}$ of G, called the *commutator* of g and h. Then $[g, h]$ is the identity if and only if g and h commute. In general, product of two commutators is not a commutator, but the set of all commutators generate a subgroup of G called the *commutator subgroup of* G, denoted by $[G, G]$. Then $[G, G]$ measures the lack of commutativity of the group G : $[G, G]$ is large for those groups G which are less commutative, and for abelian groups $[G, G]$ is the trivial subgroup $\{e\}$. If $[G, G] = G$, we may think of G as a hardcore nonabelian group, with no abelian part at all: it is known as a *perfect group.* An example of a perfect group is the group A_5 of even permutation of 5 symbols, which is a finite group of order 60. The reader is invited to prove that A_5 is perfect (hint: show first that the commutator subgroup is normal!)

What about groups which are neither abelian, nor perfect? Then $[G, G]$ is a non-trivial proper subgroup of G. We can then define a series of subgroups of G by setting $G^{(1)} = [G, G]$, and $G^{(i+1)} = [G^{(i)}, G^{(i)}]$. The series $G \supset G^{(1)} \supset G^{(2)} \supset \cdots$ is called the *derived series* of the group G. Three things can happen to it: first, some $G^{(k)}$ may happen to be a non-trivial perfect group, so that the series stops producing new subgroups after k terms and all further terms are equal to $G^{(k)}$. For such groups, the two data – the number of terms k, and the last term $G^{(k)}$ of the derived series is a measure of the degree of nonabelianness of G. Smaller the k and larger the group $G^{(k)}$, more nonabelian is the group G. Second, it is possible that some $G^{(k)}$ is the identity subgroup. In this case, we call the group G *solvable*, and we can think of G as an “almost abelian ” group. Finally, if G is infinite, it is possible that the series goes on infinitely, producing smaller and smaller groups $G^{(i)}$, but never comes to an end. Unfortunately, we do not have space here to discuss this very interesting case. The derived series therefore gives us a way of studying the degree of commutativity of groups, and it has very important applications, for example in the theory of *Lie groups*, which are groups on which one can define smooth functions and take derivatives. An example of a Lie group is the Special

Linear group $SL(n,\mathbb{R})$ of $n\times n$ matrices with determinant equal to one. We cannot leave this topic without mentioning the origin of the word *solvable* used above for groups for which $G^{(k)}=\{e\}$ for some k. The roots of the the quadratic equation $ax^2+bx+c=0$ are given by the explicit formula $\frac{1}{2a}\left(-b\pm\sqrt{b^2-4ac}\right)$. There are similar formulas for the roots of cubic and fourth-degree equations. Suppose that we want to study a polynomial p of degree n with rational coefficients, and ask: whether its roots can be represented by formulas analogous to the quadratic formula, i.e., whether the equation $p=0$ is *solvable by radicals.* To study this question, we introduce a finite group associated with the equation and called the *Galois Group* of the equation (see any algebra book, e.g. [2] for the definition.) It was discovered by the great French mathematician Éveriste Galois (1811–1832) that *a polynomial is solvable by radicals if and only if its Galois group is solvable* ! In particular, there are equations of degree five and higher whose Galois groups are not solvable, and this shows the existence of polynomial equations which are not solvable by radicals. Thus commutativity lies at the centre of the puzzle of solving polynomial equations by radicals. Galois died in a gun-fight duel over a triangular love-affair at the age of 20. But he was wise enough to make himself immortal before foolishly laying down his life.

4. Differential Calculus

Let Ω be an open set in $\mathbb{R}^n$, and let $\mathcal{C}^\infty(\Omega)$ be the infinitely many times differentiable real-valued functions on Ω. For $1\leq j\leq n$, denote by ∂_j the operator from $\mathcal{C}^\infty(\Omega)$ to itself which takes the partial derivative of a function with respect to the j-th variable: $\partial_j f=\dfrac{\partial f}{\partial x_j}$, where $(x_1,\ldots,x_n)$ denote the natural coordinates of $\mathbb{R}^n$. It is well-known that for $f\in\mathcal{C}^\infty(\Omega)$ we have $\partial_i\partial_j f=\partial_j\partial_i f$, a fact which is referred to as the *equality of mixed partials.* Another way of saying this is that the operators ∂_i and ∂_j commute under composition. If we allow the differentiation of less smooth functions, equality of mixed partials need not hold. A sufficient condition for equality of mixed partials is that all second partial derivatives of the function exist and are continuous. If the function $f:\mathbb{R}^2\to\mathbb{R}$ is given by

$$f(x_1,x_2)=\begin{cases}\dfrac{x_1x_2(x_1^2-x_2^2)}{x_1^2+x_2^2} & \text{if } x\neq 0\\ 0 & \text{if } x=0\end{cases},$$

which does not satisfy the sufficient condition stated above, we leave it to the interested reader to check that both $\partial_1\partial_2 f(0)$ and $\partial_2\partial_1 f(0)$ exist but are not equal. But we are not interested in this kind of non-commutativity, which is the result of a bad choice of f. We would like to continue to think of the ∂_j as acting on $\mathcal{C}^\infty(\Omega)$, but we want to generalize the notion of partial derivative, and ask, whether for such generalized partial derivatives, we still have commutativity. There is nothing special about the use of cartesian coordinates on Ω as far as partial derivatives are concerned. For example, if $\Omega=\mathbb{R}^2\setminus\{0\}$, we use polar coordinates (r,θ), which are related to the Cartesian coordinates (x_1,x_2) by $x_1=r\cos\theta, x_2=r\sin\theta$. A computation using the chain rule in several variables shows that for a smooth $f\in\mathcal{C}^\infty(\Omega)$ we have

$$\frac{\partial f}{\partial r}=\cos\theta\cdot\partial_1 f+\sin\theta\cdot\partial_2 f,\ \text{and}\ \frac{\partial f}{\partial\theta}=-r\sin\theta\cdot\partial_1 f+r\cos\theta\cdot\partial_2 f.$$

This suggests that as a generalization of the notion of partial derivative, we can consider operators $X:\mathcal{C}^\infty(\Omega)\to\mathcal{C}^\infty(\Omega)$ of the type $Xf=\sum_{j=1}^n a_j\cdot\partial_j f$, where $a_j\in\mathcal{C}^\infty(\Omega)$. Such an operator X is known as either a *first order linear homogeneous partial differential operator* or *a smooth vector field.* The second name comes from identifying X with the smooth vector valued function $(a_1,\ldots,a_n)$ on Ω. The partial derivative operators $\partial_j=\dfrac{\partial}{\partial x_j}$, the radial derivative operator $\dfrac{\partial}{\partial r}$ and the angular derivative operator $\dfrac{\partial}{\partial\theta}$ are all examples of smooth vector fields. Let $\mathfrak{X}(\Omega)$ denote the space of smooth vector fields on an open set Ω. We can add any two elements of $\mathfrak{X}(\Omega)$: set $(X+Y)f=Xf+Yf$, and given $a\in\mathcal{C}^\infty(\Omega)$, we can define aX by $(aX)f=a(Xf)$. Given two smooth vector fields $X,Y\in\mathfrak{X}(\Omega)$, the composition $X\circ Y$ is defined as a mapping from $\mathcal{C}^\infty(\Omega)$ to itself: $(X\circ Y)f=X(Yf)$. But $X\circ Y$ is not a smooth vector field, since a computation shows that it is a linear partial differential operator of the *second order*: if $X=\sum_{j=1}^n a_j\partial_j$ and

$Y = \sum_{k=1}^{n} b_k \partial_k$, we have for $f \in \mathcal{C}^\infty(\Omega)$ that

$$X(Yf) = \sum_{j,k=1}^{n} a_j b_k \partial_j \partial_k f + \sum_{j,k=1}^{n} a_j \partial_j b_k \partial_k f.$$

It is clear that in general $X(Yf) \neq Y(Xf)$, since the first order terms do not match (though, thanks to equality of mixed partials, the second order terms are the same!) We can measure the deviation from commutativity by constructing the *Lie bracket* $[X,Y] : \mathcal{C}^\infty(\Omega) \to \mathcal{C}^\infty(\Omega)$ defined as $[X,Y]f = X(Yf) - Y(Xf)$. But wait, the second order terms in $X(Yf)$ and $Y(Xf)$ are the same, and therefore, $[X,Y]$ only involves first order derivatives, and is actually a vector field! A calculation shows that $[X,Y] = \sum_{k=1}^{n} c_k \partial_k$, where $c_k = \sum_{j=1}^{n} (a_j \partial_j b_k - b_j \partial_j a_k)$. It is clear that on $\mathbb{R}^n$, we have $[\partial_j, \partial_k] = 0$ for each j,k and on $\mathbb{R}^2 \setminus \{0\}$ were (r,θ) are coordinates, we have $\left[\frac{\partial}{\partial r}, \frac{\partial}{\partial \theta}\right] = 0$.

We can generalize from the polar coordinates to arbitrary curvilinear coordinates $(y_1, \ldots, y_n)$ on an open set Ω. Examples of such curvilinear coordinate systems in three dimensions include the spherical and cylindrical coordinate system. By definition, the coordinates $y_1, \ldots, y_n$ are smooth functions on Ω, such that, the map $y : \Omega \to \mathbb{R}^n$ given by $y = (y_1, \ldots, y_n)$ is a smooth one-to-one onto map from Ω to an open set $U \subset \mathbb{R}^n$, and the inverse $y^{-1} : U \to \Omega$ is also a smooth map, i.e., its components are smooth. It follows that a tuple $(y_1, \ldots, y_n) \in U$ uniquely determines a point in Ω and therefore, we can use $(y_1, \ldots, y_n)$ as a new coordinate system. on Ω. Corresponding to this coordinate system, we have the partial differential operators (or smooth vector fields) $\dfrac{\partial}{\partial y_j}$, $1 \leq j \leq n$. The vector fields $\dfrac{\partial}{\partial y_j}$ are called the *coordinate vector fields* corresponding to the coordinate system $(y_1, \ldots, y_n)$. Note that the coordinate vector fields are linearly independent, since if $Z = \displaystyle\sum_{j=1}^{n} a_j \frac{\partial}{\partial y_j}$ is a linear combination of the coordinate vector fields with smooth functions $a_j \in \mathcal{C}^\infty(\Omega)$ as coefficients, we have for each k that $Zy_k = a_k$, and therefore Z is the zero vector field only if each a_k vanishes identically on Ω. Furthermore, since any smooth function may on Ω can be represented as a function of the variables $(y_1, \ldots, y_n)$, it follows that $\left[\dfrac{\partial}{\partial y_j}, \dfrac{\partial}{\partial y_j}\right]$ is identically zero, i.e., the coordinate vector fields commute.

We now come to the last example of commutativity in action in mathematics. Suppose that we are given n smooth vector fields $Y_1, \ldots, Y_n$ on an open set $\Omega \subset \mathbb{R}^n$. We can ask whether it is possible to have a curvilinear coordinate system $y_1, \ldots, y_n$ on Ω, such that $Y_j = \dfrac{\partial}{\partial y_j}$, the j-th coordinate vector field. We already have stated some necessary conditions, and it turns out they they are also sufficient, at least locally: *if $Y_1, \ldots, Y_n$ are linearly independent, and $[Y_j, Y_k] = 0$ for each j and k, then for each point p in Ω there is a curvilinear coordinate system on a neighbourgood ω of p such that on ω the Y_j are the coordinate vector fields of this curvilinear coordinate system.*

It turns out that the commutativity of families of vector fields (or more generally, the algebraic properties of their Lie brackets) is absolutely crucial in the applications of the differential calculus to geometry. Further study of this topic leads to a famous theorem called *Frobenius Theorem* (this is the same Frobenius whom we met in Section 2.) One can study more about this topic from the excellent text [3].

5. Conclusion

We have seen how commutativity and non-commutativity are important concepts in many parts of mathematics. Unfortunately, there is no space here to discuss many other interesting phenomena where this notion plays a crucial role: the notion of curvature in Riemannian geometry, the mathematical foundations of Quantum Mechanics, the notion of non-commutative geometry introduced by the great contemporary mathematician Alain Connes (born 1947, Fields Medal 1982.) If we have been able to arouse the curiosity of the reader, then our goal is fulfilled! For such a reader, we recommend the books [1,2,3] as references on the three topics discussed above.

References

[1] Ebbinghaus et. al.;*Numbers*, Graduate Texts in Mathematics, 123. Springer-Verlag, New York, 1990.
[2] N. S. Gopalakrishnan; *University Algebra (revised second edition)*, New Age International, New Delhi, 2012.
[3] S.Kumaresan; *A course in differential geometry and Lie groups*, Hindustan Book Agency, New Delhi, 2002.