Why Is It Difficult to Teach Abstract Algebra?

Man-Keung Siu
The University of Hong Kong
<mathsiu@hkucc.hku.hk>

This paper focuses on two points, relevance and abstraction, which require attention in teaching a course on abstract algebra.

Introduction

If a student and a teacher in a course on abstract algebra are to speak their mind on the course, probably they will ask each other the following questions.

Student: Why do I have to study all this abstract stuff? I do not see how it is used in my other courses or in daily life. I do not see why it is called "algebra" either. It looks entirely different from the kind of algebra I learnt in school. I get confused with the multitude of definitions and theorems about things I cannot visualize.

Teacher: All the notions are clearly defined and all the theorems are systematically derived. Everything is so precise and logical. You can even start from scratch without having to know much beforehand. How come you do not understand what is going on and feel confused?

It is not uncommon for people to regard abstract algebra as what is exemplified, but no doubt largely caricatured, by the amusing parody of A.K. Austin (*Math. Gazette*, 51 (1967), 149-150) in the form of a fictitious paper which begins with:

> "A.C. Jones in his paper "A Note on the Theory of Boffles", *Proceedings of the National Society*, 13, first defined a Biffle to be a non-definite Boffle and asked if every Biffle was reducible.
> C.D. Brown in "On a Paper by A.C. Jones", *Biffle*, 24, answered in part this question by defining a Wuffle to be a reducible Biffle and he was then able to show that all Wuffles were reducible.
> ….."

Such a distorted view of abstract algebra will naturally breed in the student doubt about and aversion to the subject. In this paper I will not elaborate on the applicability and utility of the subject. Abstract algebra is useful, as workers in the field will testify. (For example, one account based on personal experience is given in [Siu, 2000].) However, such applications are mostly beyond what an undergraduate curriculum usually encompasses. More pertinently, immediate utility is not the sole objective in justifying the teaching of a subject in the undergraduate curriculum. I would instead concentrate on the points about relevance and abstraction raised in the conversation at the beginning. There is no avoidance of these two points in a discussion on the teaching and learning of abstract algebra. I feel that difficulty in teaching is related to the lack of attention paid to these two points.

Relevance

Since I have said that I put aside the question on utility, by relevance I will not restrict the discussion to relevance with applications but will include a broader relevance with the student's past learning experience and with the historical past. Let me illustrate by a list of ten problems which I handed out on the first day of class in a course on abstract algebra which I taught for a number of years in the early 1990s. These are not meant to be answered right away, nor even by the end of the course, but are meant to bring out the relevance of the course. Some of these questions have a long history; some have played key roles in shaping the development of mathematics; some arise in applications to other fields; some can be stated in such a way as to sound familiar to a school pupil; some possess all these features. (This way in starting the course is inspired by I. Kleiner, who shared with me his teaching experience in a 1988 workshop at Kristiansand. See [Kleiner, 1995].)

(1) Why is $(-1) \times (-1) = 1$?

(2) Can one trisect an angle? duplicate a cube? square a circle?

(3) Which regular *N*-gon can be constructed?

(4) What are all integral solutions of $Y^2 = X^3 - 2$?

(5) Can one solve $2X^5 - 5X^4 + 5 = 0$ by radicals?

(6) Which is more symmetric: a square? an equilateral triangle? a rectangle? a circle?

(7) "There are a certain number of objects. If you count them by threes, two are left. If you count them by fives, three are left. If you count them by sevens, two are left. How many objects are there?" (*Sun Zi Suan Jing*, c. 4th century)

(8) Can 36 officers be drawn from 6 different ranks and from 6 different regiments so that they be arranged in a square array in which each row and each column consist of 6 officers of different ranks and different regiments? (L. Euler, 1779)

(9) How many (structural) isomers of alkanes ($C_N H_{2N+2}$) are there?

(10) Can one place N dots on an $N \times N$ grid with one dot in each row and each column so that any shifted copy has at most one dot in common in the overlapping part? (J.P. Costas, 1966; S.W. Golomb and H. Taylor, 1984)

In class I will run through these ten questions with short comments, far from sufficient to explain the question in detail (not to say the answer), but adequate as a bridge with what has been learnt before or as a preview of what lies ahead. Take Question (1) as an example. I will relate the question to students' learning experience in elementary arithmetic, how they learnt this fact and why they accepted it in school, whether they can convince someone baffled by it, or are they themselves baffled by it? (An interesting account of this learning experience is depicted in *Vie de Henri Brulard* (1836) by the French novelist Stendhal (M.-H. Beyle).) Then I will tell them about the "art of positive and negative numbers" explained in the ancient Chinese mathematical classics *Jiu Zhang Suan Shu* (Nine Chapters on the Mathematical Art, composed between 100 B.C. and A.D. 100), and the first explicit mentioning of the rule "minus times minus is equal to plus" in *Suan Xue Qi Meng* (A First Introduction to Arithmetic, 1299) by Zhu Shijie. I will continue to tell students it took the Western mathematical community a long time to come to terms with the operations on negative numbers. (I learn a lot on this topic from [Pycior, 1981].) As late as the beginning of the 19th century many English mathematicians still objected to the use of negative numbers. F. Maseres, one of the main proponents of this objection, wrote in *Tracts on the Resolution of Affected Algebraick Equations* (1800):

> "the Science of Algebra, or Universal Arithmetick, has been disgraced and rendered obscure and difficult, and disgusting to men of a just taste for accurate reasoning."

W. Frend, another of the main proponents, lamented the lack of a logical foundation for negative numbers in *The Principles of Algebra* (1796):

> "when a person cannot explain the principles of a science without reference to metaphor, the probability is, that he has never thought accurately upon the subject."

Such doubt and uneasiness still exist among pupils when they first learn the rule "minus times minus is equal to plus". In a letter to F. Maseres in 1801, R. Woodhouse, Lucasian Professor of Mathematics at Cambridge University, even said:

> "till the doctrines of negative and imaginary quantities are better taught than they are at present taught in the University of Cambridge, … they had better not be taught."

Clearly we should not adopt this policy now. We do owe our undergraduates an explanation. This motivation will naturally blend in with the discussion on solving equations related to Questions (4) and Question (5).

The discussion on solving equations is particularly relevant, not just because of the well-known fact that the word "algebra" is the Latinized version of the Arabic word "al-jabr" which appeared in the title of the book *Hisab al-jabr wàl-muqàbala* on solving equations (written by Mohammed ibn Musa al-Khowarizmi in around 825), but because it leads to subsequent development which evolve into what we today call (abstract) algebra. Classical algebra is simply the art of solving (algebraic) equations. Indeed, up to the mid 19th century mathematicians still regarded algebra as such. For a period F. Viète used the term "analysis" to denote algebra, because he did not favour the word of Arabic origin on the ground that it has no meaning in the European language! However, later in the era of I. Newton and G.W. Leibniz, calculus was regarded as an extension of algebra with lots of functions expressed as power series which behave like polynomials of infinitely high degree. The word "analysis" gradually acquired its modern meaning in describing the subject, while its original use to denote algebra never caught on. Algebra remained to be called algebra. The techniques called by F. Viète "logistica speciosa" that deals with operation on symbols representing species or forms of things, as contrasted to "logistica numerosa" that deals with the arithmetic of numbers, are still being learnt in school algebra in order to prepare the way for solving equations. Unfortunately, a central message is normally lost in school algebra amidst the technicalities of simplification of algebraic expressions, factorization of polynomials, laws of indices, etc. This central message is:

> We treat numerical quantities as *general objects* and manipulate such general objects as if they are numerical quantities. Although we do not know what they are (prior to solving the equation) we know they stand for certain numbers, and as such they obey *general rules* (such as $a + (-a) = 0$, $a \times (b \times c) = (a \times b) \times c$, etc.) We can therefore apply these general rules systematically to solve problems which can be formulated in terms of equations.

No wonder F. Viète was so optimistic as to close his book *In Artem Analyticam Isogoge* (Introduction to the Analytical Art, 1591) with the saying "Quod est, Nullum non problema solvere" (there is no problem that cannot be solved)! This central message continues to ring true in the subsequent development of algebra. In [MacLane 1981] S. MacLane describes abstract algebra as:

> "the program of studying algebraic manipulations on arbitrary objects with the intent of obtaining theorems and results deep enough to give substantial information about the prior existing particular objects."

A study of this kind is facilitated by, if not necessitates, the use of the axiomatic method, but one must not equate abstract algebra with axiomatic approach. After all, the latter is only an organizing principle but not the substance of the former. We will treat this point in more detail when we discuss abstraction in the next section.

Coming back to solving equations, I agree that the full story cannot be understood until we cover Galois theory, which is not usually covered, for lack of time and for the level of sophistication, in a first course in abstract algebra. But some indication of how this problem is related to the emergence of the concept of a group is worth the time spent even in a first course in abstract algebra. From the symmetry of an equation to the symmetry in geometry and beyond is a trip students can enjoy and benefit from the course. To illustrate I will now give an extremely sketchy explanation on solving the cubic equation $X^3 + bX^2 + cX + d = 0$, only highlighting the crucial point on symmetry. (This explanation was given by A.-T. Vandermonde in a brilliant memoir of 1774 [Tignol, 1980].) Let α, β, γ be the three roots of the cubic equation. Note that

$$\alpha = \frac{1}{3}\left[(\alpha+\beta+\gamma) + \sqrt[3]{(\alpha+\omega\beta+\omega^2\gamma)^3} + \sqrt[3]{(\alpha+\omega^2\beta+\omega\gamma)^3}\right]$$

where $\omega = e^{2\pi i/3}$. Similar expressions can be found for β and γ. If we put $u = \alpha+\omega\beta+\omega^2\gamma$ and $v = \alpha+\omega^2\beta+\omega\gamma$, then $\alpha = \frac{1}{3}\left[-b + \sqrt[3]{u^3} + \sqrt[3]{v^3}\right]$. Can we express u^3, v^3 in terms of the coefficients of the original equation (and constants) using only rational operations and radicals? If we can, then we can express all roots α, β, γ in terms of the coefficients of the original equation (and constants) using only rational operations and radicals, i.e. the cubic equation is solvable by radicals. Note that $u^3 = \frac{1}{2}\left[u^3 + v^3 + \sqrt{(u^3 - v^3)^2}\right], v^3 = \frac{1}{2}\left[u^3 + v^3 - \sqrt{(u^3 - v^3)^2}\right]$ and $(u^3 - v^3)^2 = (u^3 + v^3)^2 - 4u^3v^3$. By a result on symmetric polynomials (which is familiar to school pupils in the special case of polynomials in two indeterminates), we will achieve that if we can show that $u^3 + v^3$ and u^3v^3 are symmetric polynomials in α, β, γ. How can we check this? Let us run through all six possible permutations of α, β, γ and see what $u = \alpha+\omega\beta+\omega^2\gamma$ and $v = \alpha+\omega^2\beta+\omega\gamma$ become? For instance, if α becomes α, β become γ and γ becomes β, then u becomes v and v becomes u so that u^3 becomes v^3 and v^3 becomes u^3. Take one more case, if α becomes γ, β becomes α and γ becomes β, then u becomes ωu and v becomes $\omega^2 v$ so that u^3 becomes u^3 and v^3 becomes v^3. A simple calculation will show that for all six permutations of α, β, γ, $u^3 + v^3$ remains $u^3 + v^3$ and u^3v^3 remains u^3v^3. Hence, $u^3 + v^3$ and u^3v^3 are symmetric polynomials in α, β, γ. More generally, the set of certain permutations on a finite number of symbols under the operation of composition is a typical example of a group. The notion of a group plays a prominent role in the discussion on the problem of solving equations in radicals.

Abstraction

In the last section I touch upon the point on abstraction. Let us listen to what L. Kronecker had to say in his 1870 paper on algebraic number theory:

> "these principles belong to a more general and more abstract realm of ideas. It is therefore proper to free their development from all inessential restrictions, thus making it unnecessary to repeat the same argument when applying it in different cases … Also, when stated with all admissible generality, the presentation gains in simplicity and, since only the truly essential features are thrown into relief, in transparency."

That much is well said, but only for the *teachers* who are seasoned mathematicians themselves. For *students*, especially those who first embark on higher mathematics, that can only give them, if anything at all, a comforting psychological support for what is to come. In reality that is not enough to arm them to face the challenge. In its long process of evolution (which is unfortunately unfamiliar to most students) mathematics has acquired a language of its own, which can sound quite obscure to one not steeped in that training. Abstraction is a forte that lends mathematics its power, though it causes anxiety in many as well. However, we should see that as a challenge rather than as something to avoid, to uncover the concrete parts which evolve into the abstract concepts rather than to only study the concrete parts.

Having stated the generalities above I will now focus on three specific aspects pertaining to abstract algebra: (i) definitions, (ii) proofs, (iii) symbolic thinking? geometric thinking? or something else?

For mathematicians a definition possesses two levels in meaning. On the first level, a definition serves the mere purpose of an abbreviation. Instead of saying:

> If straight lines ℓ_1, ℓ_2, L on the same plane are such that ℓ_1 sets up on L adjacent angles equal to one another and ℓ_2 sets up on L adjacent angles equal to one another, then ℓ_1 and ℓ_2 do not meet one another in either direction when they are produced indefinitely,

we need only say:

> If ℓ_1, ℓ_2 are perpendicular to L, then ℓ_1, ℓ_2 are parallel.

On the second level, a definition embodies a concept. The job of a mathematician includes that of formulating useful definitions and delineating relationship between definitions. (When we retreat to more and more basic definitions we come face to face with the role of axioms.) Only because we understand a definition on the second level well enough do we feel easy about it on the first level, just like what Shakespeare says, "What's in a name? That which we call a rose by any other name would smell as sweet." (*Romeo and Juliet*, *Act 2*, *Scene 1*) For a teacher, clarity and preciseness are all that matter. But for a novice, if not sufficient attention is given to a definition, they will regard a definition as something coming out of the blue, something mysterious and incomprehensible, and hence something to be memorized in order to pass the examination. At this point, communication between the teacher and the students already breaks down. In a paper on mathematical definitions H. Poincaré says (*L'Enseignement Mathématique, 6 (1904), 255-283*):

> "What is a good definition? For the philosopher or the scientist, it is a definition which applies to all the objects to be defined, and applies only to them, … in education it is not that; it is one that can be understood by the pupils, … How are we to find a statement that will at the same time satisfy the inexorable laws of logic and our desire to understand the new notion's place in the general scheme of the science, our need of thinking in images? More often than not we shall not find it, and that is why the statement of a definition is not enough; it must be prepared and it must be justified."

I will illustrate with the definition of a quotient structure, which is a notorious learning difficulty for an average undergraduate. At the same time it is a notion which appears in many contexts and warrants the time and effort for its explication. In elementary number theory it appears in the form of modulo arithmetic, which can be traced back to the work of C.F. Gauss in 1801. In the theory of system of linear equations (respectively *n*th order linear recurrence relation, respectively *n*th order linear differential equation), it appears in the form of the quotient space of a suitable vector space modulo the solution space of the associated homogeneous system. In the theory of groups it appears in the form of the quotient group modulo a normal subgroup. In topology it appears in the form of a quotient topological space modulo a subspace. Historically speaking, the notion made its first explicit début in the explanation by A.L. Cauchy on what the field of complex numbers is, viz the quotient ring of polynomials with real coefficients modulo the ideal generated by $X^2 + 1$. Although the contexts are different and the purposes of making use of the quotient structure may vary, there is an underlying common principle, the partition of a set through the identification of certain elements of the set. A partition of a set amounts to the same thing as the introduction of an equivalence relation on the set. That explains both the motivation why we do that and the technique on how to do that. It also reveals the main learning obstacle in that we are looking at the process (equivalent relation) and the object (equivalence class) at the same time. Worse yet, we have to learn how to see a subset of elements (coset) as an element by itself without losing sight that it actually stands for a subset of elements. It is this kind of flexibility of framework which is demanding on the mathematical maturity of the students. It takes time to let the idea sink in. Dishing out a correct and precise definition of, say a quotient group, is not enough, although it sounds perfectly clear to the teacher, who will wonder why students cannot take in such a clear-cut answer.

Let us continue to see the proof of one theorem making use of a quotient structure. As Y.I. Manin puts it, "a good proof is what makes us wiser". Our aim is to explain and to persuade, not just to verify and to force the result upon the learner. The theorem we will look at is a most basic result in the theory of finite groups usually referred to as Lagrange's Theorem (J.L. Lagrange 1770/1771):

> If H is a subgroup of a finite group G, then the order of H is a divisor of the order of G.

How do we visualize the result? A traditional way to classify mathematical thinking is to say that some people are more inclined towards geometric thinking and some are more inclined towards symbolic thinking. Even in abstract algebra both types of thinking can be useful. However, in abstract algebra there may be a third type in which a schematic diagram aids the thinking. For instance, in this example one tries to figure out a way to count the elements of a group *G*. Part of *G* is the subgroup $H = \{e = h_1, h_2, \ldots, h_s\}$ where $s = |H|$. If *H* is the whole of *G*, then we are done. If not, then there is some *g* in *G* outside of *H*. It is not hard to see that we create another (disjoint from *H*) subset $gH = \{g = gh_1, gh_2, \ldots, gh_s\}$. (It needs some checking for the qualification "another".) If *H* and *gH* together cover the whole of *G*, then again we are done. If not, then there is still some *g'* in *G* which is outside of *H* and *gH*. (Some doodling on the paper may help!) Consider $g'H = \{g' = g'h_1, g'h_2, \ldots, g'h_s\}$, which turns out to be yet another (disjoint from *H* and *gH*) subset. (It needs even more checking for the qualification "yet another".) If *H* and *gH* and *g'H* together cover the whole of *G*, then we are done. If not, we repeat the process until we arrive at a partition of *G* into *t* pieces. Now that we realize the connection between a partition and an equivalence relation, we can streamline the proof of Lagrange's Theorem by passing to the quotient set *G/H* of cosets of *G* by *H*. These *t* copies of H, each consisting of $s = |H|$ elements, exhaust the whole of *G* with $|G|$ elements, so $t|H| = |G|$, i.e. $|H|$ is a divisor of $|G|$. A further natural question is to ask whether we can turn *G/H* into a group by inducing the group operation of *G* on *G/H*. It turns out this is not always possible, but will be possible if and only if *H* is what we term a normal subgroup of *G*. That will lead to further discussion on why we want to do so. My favourite examples to illustrate why we want to "kill" a subgroup are: (i) to look at the commutator subgroup [*G*, *G*] and the quotient group *G*/[*G*, *G*] to discuss "how abelian" *G* is? (ii) to look at the centre *Z* and the quotient group *G*/*Z* to discuss when *G* is abelian. In this way the lessons go on and the course unfolds.

References

Kleiner, I. (1995). The teaching of abstract algebra: An historical perspective. In F. Swetz, J. Fauvel, O. Bekken, B. Johansson & V. Katz (Eds.) *Learn From the Masters*! Washington D.C.: Mathematical Association of America.

MacLane, S. (1981). History of abstract algebra: Origin, rise, and decline of a movement. *Texas Tech. University Mathematics Series*, 13, 3-35.

Pycior, H.M. (1981). George Peacock and the British origins of symbolical algebra, *Hist. Math.*, 8, 23-45.

Siu, M.K. (2000). Combinatorics and algebra: A medley of problems? A medley of techniques. *Contemporary Mathematics*, 264, 287-305.

Tignol, J.-P. (1980). *Leçons sur la théorie des équations*. Louvain-la-Neuve: Institut de Mathématique Pure et Appliquée, Université Catholique de Louvain.

Why do we teach math? (Image source). It could be because the mathematical procedures that are taught in schools will be useful to students later, but I am pretty sure this is false. Almost everyone forgets those procedures as they get older because most people in our society use virtually none of the procedures they learned in school in their day-to-day life.Â We could also ask students to talk about mathematics in the abstract and come to a shared understanding of what elegance and beauty in mathematics mean. As far as I know, none of these activities is a common one in math classes.Â I think that we teach algebra because it is a bridge between the more concrete ideas children have developed about numbers, into the more abstract realm of ideas about classes of numbers. Even though abstract algebra can be difficult, there are some things that can make it a less difficult class. As in most math classes, the topics in abstract algebra build on themselves. If you can put a consistent effort into the class right from the start of the semester, making sure that you understand each theorem and definition as you go, and you are reasonably comfortable with proofs, you should be able to do well in the class. However, if you allow yourself to fall behind, early on, you will likely have a hard time in the class. This is why it is very important to make sure to dedicate "Why study algebra?" If you're a parent, it's a question that you will no doubt hear as your children study the subject. If you're a student, it is a very natural question to ask, "What's the point of learning algebra in the first place?" After all, all of the math leading up to algebra that we learned growing up such as addition, multiplication, decimals, fractions, and the like, seem to have a concrete meaning. These concepts all deal with numbers in some way or another and because of this we can wrap our brains more easily around the concepts.Â This is a difficult question, but the simplest answer is that Algebra is the beginning of a journey that gives you the skills to solve more complex problems. What types of problems can you solve using only the skills you learned in Algebra?