VICTOR J. KATZ

STAGES IN THE HISTORY OF ALGEBRA WITH IMPLICATIONS FOR TEACHING*

ABSTRACT. In this article, we take a rapid journey through the history of algebra, noting the important developments and reflecting on the importance of this history in the teaching of algebra in secondary school or university. Frequently, algebra is considered to have three stages in its historical development: the rhetorical stage, the syncopated stage, and the symbolic stage. But besides these three stages of expressing algebraic ideas, there are four more conceptual stages which have happened along side of these changes in expressions. These stages are the geometric stage, where most of the concepts of algebra are geometric ones; the static equation-solving stage, where the goal is to find numbers satisfying certain relationships; the dynamic function stage, where motion seems to be an underlying idea, and finally, the abstract stage, where mathematical structure plays the central role. The stages of algebra are, of course not entirely disjoint from one another; there is always some overlap. We discuss here high points of the development of these stages and reflect on the use of these historical stages in the teaching of algebra.

KEY WORDS: algebra, abstract stage, equation-solving stage, function stage, geometric stage, rhetorical stage, symbolic stage, syncopated stage

This article is about algebra. So the first question is, what do we mean by algebra? Few secondary textbooks these days give a definition of the subject, but that was not the case two centuries ago. For example, Colin Maclaurin wrote, in his 1748 algebra text, “Algebra is a general Method of Computation by certain Signs and Symbols which have been contrived for this Purpose, and found convenient. It is called an Universal Arithmetic, and proceeds by Operations and Rules similar to those in Common Arithmetic, founded upon the same Principles.” (Maclaurin, 1748, p. 1) Leonhard Euler, in his own algebra text of 1770 wrote, “Algebra has been defined, *The science which teaches how to determine unknown quantities by means of those that are known.*” (Euler, 1984, p. 186) That is, in the 18th century, algebra dealt with determining unknowns by using signs and symbols and certain well-defined methods of manipulation of these. Is that what algebra still is? It is hard to say. Sometimes, algebra is defined as “generalized arithmetic,” whatever that means. But a look at a typical secondary algebra textbook reveals a wide variety of topics. These include the arithmetic of signed numbers, solutions of linear equations, quadratic

*Commentary from a Mathematics Educator Bill Barton. See also the last page.

Educational Studies in Mathematics (2007) 66: 185–201
DOI: 10.1007/s10649-006-9023-7

equations, and systems of linear and/or quadratic equations, and the manipulation of polynomials, including factoring and rules of exponents. The text might also cover matrices, functions and graphs, conic sections, and other topics. And, of course, if you go to an abstract algebra text, you will find many other topics, including groups, rings, and fields. So evidently, "algebra" today covers a lot of ground.

In this article, we will take a rapid tour of the history of algebra, first as it was defined in the eighteenth century and then as it is understood today. We want to look at where algebra came from and why? What was its original purpose? How were the ideas expressed? And how did it get to where it is today? Finally, we want to consider how the history of algebra might have some implications for the teaching of algebra, at whatever level this is done.

In many history texts, algebra is considered to have three stages in its historical development: the rhetorical stage, the syncopated stage, and the symbolic stage. By the rhetorical, we mean the stage where are all statements and arguments are made in words and sentences. In the syncopated stage, some abbreviations are used when dealing with algebraic expressions. And finally, in the symbolic stage, there is total symbolization – all numbers, operations, relationships are expressed through a set of easily recognized symbols, and manipulations on the symbols take place according to well-understood rules.

These three stages are certainly one way of looking at the history of algebra. But I want to argue that, besides these three stages of expressing algebraic ideas, there are four conceptual stages that have happened along side of these changes in expressions. The conceptual stages are the geometric stage, where most of the concepts of algebra are geometric; the static equation-solving stage, where the goal is to find numbers satisfying certain relationships; the dynamic function stage, where motion seems to be an underlying idea; and finally the abstract stage, where structure is the goal. Naturally, neither these stages nor the earlier three are disjoint from one another; there is always some overlap. I will consider both of these sets of stages to see how they are sometimes independent of one another and at other times work together. But because the first set of stages is well known and discussed in detail by Luis Puig in the recent *ICMI Study on Algebra* (Puig, 2004), I will concentrate on the conceptual ones.

We begin at the beginning of algebra, whatever it is. It would seem that the earliest algebra – ideas which relate to the eighteenth century definitions by Euler and Maclaurin – comes from Mesopotamia starting about 4000 years ago. Mesopotamian mathematics (often called Babylonian mathematics) had two roots – one is accountancy problems, which from the beginning were an important part of the bureaucratic system of the earliest Mesopotamian dynasties, and the second is a "cut and paste" geometry probably developed by surveyors as they figured out ways to understand the division of land. It is chiefly out of this cut and paste geometry that what we call Babylonian algebra grew. In particular, many old-Babylonian clay tablets dating from 2000–1700 BCE contain extensive lists of what we now call quadratic problems, where the goal was to find such geometric quantities as the length and width of a rectangle. In accomplishing this goal, the scribes made full use of the surveyors' "cut-and-paste" geometry.

As an example, we consider the problem from tablet YBC 4663 (c. 1800 BCE), in which we are given that the sum of the length and width of a rectangle is 6 1/2, and the area of the rectangle is 7 1/2. (Neugebauer and Sachs, 1945, p. 70) We are to find the length and the width. The scribe describes in detail the steps he goes through. He first halves 6 1/2 to get 3 1/4. Next, he squares 3 1/4 to get 10 9/16. From this (area), he subtracts the given area 7 1/2, giving 3 1/16. The square root of this number is extracted – 1 3/4. Finally, the scribe notes that the length is 3 1/4 + 1 3/4 = 5, while the width is 3 1/4 – 1 3/4 = 1 1/2. It is quite clear that the scribe is dealing with a geometric procedure. In fact, a close reading of the wording of the tablets seems to indicate that the scribe had in mind this picture, where for the sake of generality (since there are lots of similar problems solved the same way) the sides have been labeled in accordance with the generic system which we write today as $x + y = b$, $xy = c$ (Figure 1).

The scribe began by halving the sum b and then constructing the square on it. Since $b/2 = x - ((x - y)/2) = y + ((x - y)/2)$, the square on $b/2$ exceeds the original rectangle of area c by the square on $(x - y)/2$, that is, $((x + y)/2)^2 - xy = ((x - y)/2)^2$. The figure then shows that if one adds the side of this latter square to $b/2$, one finds the length x, while if one subtracts it from $b/2$, one gets the width y. We can express the algorithm as a modern-day formula:

$$x = \frac{b}{2} + \sqrt{\left(\frac{b}{2}\right)^2 - c} \quad y = \frac{b}{2} - \sqrt{\left(\frac{b}{2}\right)^2 - c}$$

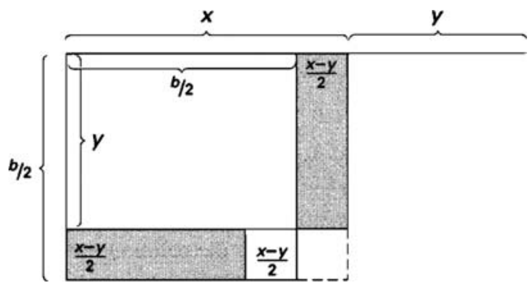

Figure 1.

I emphasize that this is a modern formula. The Babylonians had nothing similar on their tablets. What they described, totally in words, was a procedure, an algorithm. I have just translated this for our convenience – because we are accustomed to doing it this way. But the Babylonians were certainly in the "rhetorical" stage, the first stage mentioned earlier.

So, with this problem and other similar problems, expressed in words yet solved using geometric ideas, we have what I call the beginning of algebra, the beginning of a process of solving numerical problems via manipulation of the original data according to fixed rules. This example is only one of many found on clay tablets asking to determine geometric quantities, so somehow dealing with those matters was the object of this first algebra.

In Greece, of course, mathematics was geometry. Yet what we think of as algebraic notions were certainly present in the work of Euclid and Apollonius. There are numerous propositions, particularly in Book II of Euclid's *Elements* (Euclid, 2002) that show how to manipulate directly rectangles and squares. And then there are propositions where Euclid "solves" what appear to be algebraic problems for geometric results, such as the position of a particular point on a line. Euclid solved these problems by manipulating geometric figures, but, unlike the Babylonians, he based the manipulations on clearly stated axioms.

Let us look at *Elements* Proposition II-5: *If a straight line is cut into equal and unequal segments, the rectangle contained by the unequal segments of the whole together with the square on the straight line between the points of section is equal to the square on the half.*

If we think of the "unequal segments of the whole" as x and y (with sum b), then the proposition seems to say that $xy + (\frac{x-y}{2})^2 = (\frac{x+y}{2})^2$ and this result can be used to solve the system of equations $x + y = b; xy = c$. If we substitute c for xy and b for $x + y$, we get

$$\left(\frac{x-y}{2}\right)^2 = \left(\frac{b}{2}\right)^2 - c \quad \text{or} \quad \frac{x-y}{2} = \sqrt{\left(\frac{b}{2}\right)^2 - c}.$$

Then

$$x = \frac{x+y}{2} + \frac{x-y}{2} = \frac{b}{2} + \sqrt{\left(\frac{b}{2}\right)^2 - c},$$

with a similar result for y.

Proposition II-5 does not itself solve an equation. Islamic mathematicians, however, quoted this result centuries later to justify their own algorithmic solution of quadratic equations. However, Euclid himself does solve what we can call equations. Some of these are in Book VI of the *Elements*, but they are much clearer in another work, the *Data* (Euclid, 1993). Consider

Proposition 1 *If two straight lines contain a given area in a given angle, and if the sum of them be given, then shall each of them be given (i.e., determined).*

If we take the given angle to be a right angle - and the diagrams in the surviving medieval manuscripts show such an angle - then the problem is essentially identical to the standard Babylonian problem of finding the sides of a rectangle given the area and the semi-perimeter. In fact, Euclid's method is also virtually the same as that of the Babylonians. To demonstrate this proposition, Euclid sets up a rectangle, one of whose sides is $x = AS$, the other $y = AC$. He then draws $BS = AC$ and completes the rectangle $ACBD$. Now $AB = x + y$ is given as is the area of rectangle $ACFS$ (Figure 2). So to determine AS and AC, he must apply

Proposition 2 *If a given area be applied to a given straight line, falling short by a figure given in species, the sides of the deficiency are given.*

Here the given area is that of $ACFS$ and this has been applied to the given line AB, but it falls short by a square. Euclid claims that he can determine the sides of the square, namely y. To do this, he bisects AB at E, constructs the square on BE, then notes that this square is equal to the sum of the rectangles $ACFS$ and the little square at the bottom (Figure 3).

Since the rectangle's area is given, say c, and since the area of the square is given, $(b/2)^2$, the little square at the bottom is the difference of c and $(b/2)^2$. We can see that y, the "sides of the deficiency" is known, and so is

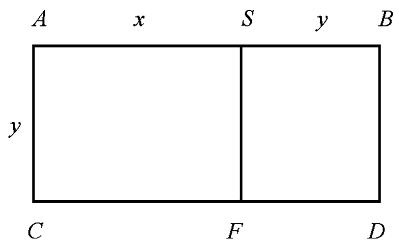

Figure 2.

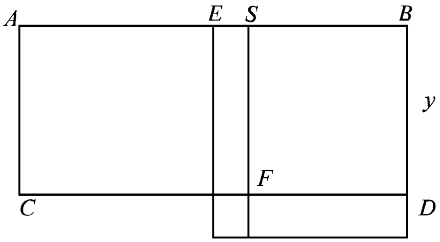

Figure 3.

x, the length of the given area. In fact, we have

$$x = \frac{b}{2} + \sqrt{\left(\frac{b}{2}\right)^2 - c} \quad y = \frac{b}{2} - \sqrt{\left(\frac{b}{2}\right)^2 - c}$$

the standard Babylonian formula for this case.

There are many parallels between Babylonian algebra expressed geometrically and Greek "geometric algebra," so a natural question to ask is whether the Greek material is an adaptation of material that the Greeks learned from the Babylonians. There are arguments supporting both sides of this question, but the answer is still unknown. Whether or not there was transmission, it is clear that the "algebra" we recognize in the Greek geometrical works is based on geometric manipulation, just as the algebra in the Babylonian tablets.

Of course, even though the underlying rationale for Babylonian equation solving was geometric, the Babylonians still developed algorithms, or procedures, to solve equations. Eventually, the algorithms began to replace the geometry. The history of algebra begins moving to the "equation solving" stage. We see evidence of this in Diophantus's knowledge of the algorithm for solving quadratic equations, solely based on numbers, in the third century. In India, the quadratic formula also appears without any geometric underpinning as early as the sixth century.

The first true algebra text which is still extant is the work on *al-jabr* and *al-muqabala* by Mohammad ibn Musa al-Khwarizmi, written in Baghdad around 825 (Al-Khwarizmi, 1831). The first part of this book is a manual for solving linear and quadratic equations. Al-Khwarizmi classifies equations into six types, three of which are mixed quadratic equations. For each type, he presents an algorithm for its solution. For example, to solve the quadratic equation of the type "squares and numbers equal to roots" ($x^2 + c = bx$), al-Khwarizmi tells his readers to take half the number of "things", square it, subtract the constant, find the square root, and then add it to or subtract it from the half the roots already found. As in Babylonian times 28 centuries earlier, the algorithm is entirely verbal. There are no symbols.

Having written down an algorithm, al-Khwarizmi justifies it using a "cut-and-paste" geometry, very much like the Babylonians. But once the justifications are dispensed with, al-Khwarizmi only expects the reader to use the appropriate algorithm. This is different from the Babylonian procedure, in which each problem indicates some use of the geometric background. In another difference with his Babylonian predecessors, al-Khwarizmi virtually always presents abstract problems, rather than problems dealing with lengths and widths. Most of the problems, in fact, are similar to this one: "I have divided ten into two parts, and having multiplied each part by itself, I have put them together, and have added to them the difference of the two parts previously to their multiplication, and the amount of all this is fifty-four." The equation translating this problem is $(10 - x)^2 + x^2 + (10 - x) - x = 54$. Al-Khwarizmi reduces this to $x^2 + 28 = 11x$ and then solves according to his algorithm.

Al-Khwarizmi does, however, have one or two other types of problems: "You divide one dirhem among a certain number of men, which number is 'thing.' Now you add one man more to them, and divide again one dirhem among them; the quota of each is then one-sixth of a dirhem less than at the first time." Al-Khwarizmi describes how to translate this problem into the equation $x^2 + x = 6$; he can then use one of his algorithms to find that $x = 2$.

Algebra has now moved decisively from the original geometric stage to the static equation-solving stage. Al-Khwarizmi wants to solve equations. And an equation has one or two numerical answers. His successors in the Islamic world do much the same thing. They set up quadratic equations to solve and then solve them by an algorithm to get one or two answers. You may notice that I am only talking here about quadratic equations. Surely, Islamic mathematicians solved linear equations. In his *al-jabr*, Al-Khwarizmi has numerous problems solvable by linear equations, mostly in his section on inheritance problems. But to a large extent, solving linear equations was part of what we would call arithmetic, not algebra. That is, the basic ideas were part of proportion theory, an arithmetical concept.

Over the next few centuries, Islamic mathematicians worked out various ideas in algebra. They developed all the procedures of polynomial algebra, including the rules of exponents, both positive and negative, and the procedures for dividing as well as multiplying polynomials. Yet the goal of these manipulations was to solve equations, and since the Islamic mathematicians could not solve equations of degree higher than two by an algorithm, they developed two alternative methods. First, there was a return to geometry, but a more sophisticated geometry than Euclid's. Namely, Omar Khayyam found a way to solve cubic equations by determining the intersection of particular conic sections. A second alternative, and one that was certainly more useful, was to determine numerical ways of approximating the solution, ways closely related to what has become known as the Horner method. Still, of course, the idea was to find a single answer (or maybe two or three).

One Islamic mathematician, who was interested in solving cubic equations, began to think in new ways. This was Sharaf al-Din al-Tusi (d. 1213), a mathematician born in Tus, Persia. Let us consider his analysis of the equation $x^3 + d = bx^2$. He began by putting the equation into the form $x^2(b - x) = d$. He then noted that the question of whether the equation has a solution depends on whether the "function" on the left side reaches the value d or not. To determine this, he needed to find a maximum value for the function. Although he does not tell us how he did so, he claims and then proves that the maximum value occurs when $x = 2b/3$, which in fact gives the functional value $4b^3/27$. Thus Sharaf al-Din could now claim that if this value is less than d, there are no (positive) solutions; if it is equal to d, there is one solution at $x = 2b/3$, and if it is greater than d, there are two solutions, one between 0 and $2b/3$ and one between $2b/3$ and b (Figure 4).

Sharaf al-Din still could not figure out an algorithm to determine these solutions, but at least he knew the basic conditions on whether the solutions existed. Unfortunately, his work was not followed up, either in Islam or later in Europe. So this attempt in Islam to move to "functions" ultimately got nowhere. One of the reasons, perhaps, is that Sharaf used no symbols – and dealing with functions without symbols is very difficult.

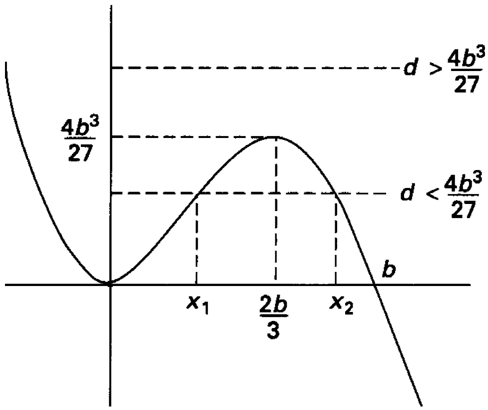

Figure 4.

The Islamic algebra which was transmitted to Europe in the twelfth and thirteenth centuries was just the static equation-solving algebra. There were several routes that al-Khwarizmi's algebra took into Europe, including the work of Leonardo of Pisa (Fibonacci), and Abraham bar Hiyya (in Spain – d. 1136), as well as the direct translations made by Robert of Chester and Gerard of Cremona. In all of them, the basic idea of static equation – solving remained. We have a problem to solve, in general a numerical problem, which involves squares. We figure out which algorithm to use and then use it to get our answer (or in some cases, two answers). Interestingly, the problems considered only involved positive numbers. Negative numbers did not appear, either in the problem posed or in the solution. There was never a perceived need to solve an equation such as $x + 3 = 2$, nor a "real" problem that required such an equation.

In all the algebra texts that reached Europe and in the ones that Europeans wrote soon thereafter, the problems solved were virtually always abstract problems. It was very difficult to come up with a quadratic equation problem coming from the "real world". One mathematician who did so was the sixteenth century English mathematician Robert Recorde. Here is one of his problems: There is a strange journey appointed to a man. The first day he must go 1 1/2 miles, and every day after the first he must increase his journey by 1/6 of a mile, so that his journey shall proceed by an arithmetical progression. And he has to travel for his whole journey 2955 miles. In what number of days will he end his journey? (Recorde, 1969) Of course, Recorde is assuming you know how to sum an arithmetic progression. And it is out of this sum that the quadratic equation emerges.

In sixteenth century Italy, there was a major breakthrough in mathematics. Several Italian mathematicians figured out how to solve cubic equations and fourth degree equations as well. But as we see in Girolamo Cardano's *Ars Magna* (Cardano, 1968), the basic principle here was the same as in the solution of quadratic equations. First, Cardano classified cubic equations into a large number of different classes. For each class, he presented an algorithm for solution. He often justified the algorithms by some geometric argument, but basically it was the algorithm itself that was important. It is that which enabled one to find an answer (or perhaps two or three) to a very clearly defined equation. But that was it. Algebra was still just about finding solutions to equations. And if you look through Cardano's work, you find that again virtually all the problems he solves by using equations are purely abstract.

However, beginning with Cardano and the other Italian algebraists of his time, we begin to see a very rapid change from the Islamic rhetorical algebra through the syncopated stage into the modern symbolic stage. All are, however, in the context of static equation solving. The easiest way to comprehend the change is to look at the notation by which four mathematicians, Cardano, Francois Viète, Thomas Harriot, and René Descartes expressed Cardano's algorithm for solving the cubic equation "cube and thing equal to number," that is, $x^3 + cx = d$. First, here is Cardano's solution to the equation $x^3 + 6x = 20$. (Cardano had no way of expressing a "general" equation.)

$$\Re\,\mathrm{v} : \mathrm{cub} : \Re 108\ \mathrm{p} : 10\,\mathrm{m} : \Re\,\mathrm{v} : \mathrm{cubica}\,\Re 108\,\mathrm{m}{:}10.$$

Next, here is Viète's solution to A cube + B plane 3 in A equals Z solid 2:

$$A \text{ is } l.c.l.\overline{\overline{B\text{plane plane plane} + Z\text{solid solid}} + Z\text{solid}} - l.c.l.\overline{\overline{B\text{plane plane plane} + Z\text{solid solid}} - Z\text{solid}}$$

Then, we have Harriot's solution to $2ccc = 3bba + aaa$:

$$a = \sqrt[3]{\sqrt{bbbbbb + cccccc} + ccc} - \sqrt[3]{\sqrt{bbbbbb + cccccc} - ccc}$$

It is only a short step from Harriot's notation to that of Descartes. Here is Descartes' solution to the equation $z^3 = -pz + q$:

$$\sqrt{\mathrm{C.} + \frac{1}{2}q + \sqrt{\frac{1}{4}qq + \frac{1}{27}p^3}} - \sqrt{\mathrm{C.} - \frac{1}{2}q + \sqrt{\frac{1}{4}qq + \frac{1}{27}p^3}}$$

With a new notation coming into place in the seventeenth century, a great change in point of view was also taking place in algebra itself. Mathematicians started asking questions other than "find the solution to that problem expressed as an equation." There are probably many reasons for this, but certainly one of the reasons was increasing interest in astronomy and physics. Johann Kepler was interested in the path of the planets. Galileo Galilei was interested in the path of a projectile. In both of these cases, it was not a "number" that was wanted, but an entire curve. Both Kepler and Galileo realized that the solutions to their problems were conic sections, and the only way they knew how to deal with these was by what they had learned from Apollonius. His mathematics was largely "static" in that he was not concerned with moving points – just with a particular slice of a cone. Nevertheless, Kepler and Galileo were able to pull out of his work the ideas they needed to represent motion.

Now neither Kepler nor Galileo had a useful notation for representing motion. They did not use algebra, but relied on Greek models, including the detailed use of proportion theory. For example, Galileo, in his *Two New* Sciences of 1638, describes the path of a moveable object projected on a horizontal plane which eventually ends. He writes: "the moveable, ... driven to the end of this plane and going on further, adds on to its previous equable and indelible motion that downward tendency which it has from its own heaviness. Thus there emerges a certain motion, compounded from equable horizontal and from naturally accelerated downward motion, which I call projection." (Galileo, 1974, p. 217) And then he demonstrates the following, by the use of a purely geometrical argument, with no algebraic symbolism at all:

Proposition 3 *When a projectile is carried in motion compounded from equable horizontal and from naturally accelerated downward motions, it describes a semiparabolic line in its movement.*

If one reads Kepler's *New Astronomy* of 1609 (Kepler, 1992), one finds similar ideas. Besides going through all of the numerical data from the observations, Kepler discusses in great detail the geometry of the ellipse as he finally shows that the orbit of Mars is an ellipse. There is no algebra in the book, only geometry. This book was very difficult to read. One had to fight through all of the verbiage – and Kepler used lots of it. It seemed that something had to be done so that the important physical results of Kepler and Galileo could be better disseminated and so that further developments could result from their work.

In 1637, the appropriate tools for representing this work appeared. The two fathers of analytic geometry, Fermat and Descartes, produced their first works on the subject. Both were interested in the use of algebra to represent curves, although, interestingly, neither cited any motivations from physics. Descartes' clear motivation was to use the algebra to solve geometric problems, while Fermat was just interested in representing curves through algebra. But since both showed how to represent a curve, however described verbally, through algebra, analytic geometry gave mathematicians a mechanism for representing motion. And Newton, for one, picked up on this as he developed the calculus.

Curiously, Newton, when he wrote the *Principia*, was somewhat hesitant to use algebra. Much of the work is presented using classical geometry, although with a modern "twist." But given the newly developed algebra, which Newton certainly used freely in some of his other works, it is not surprising that one of the major thrusts in the early eighteenth century was to translate Newton's ideas into algebraic language and prove them using algebra and the newly invented calculus. The mathematicians who accomplished this, along with others in that time period, were no longer interested in finding a "number" as answer to a problem. They wanted a curve. They were interested in seeing how objects – be they planets or projectiles - moved, and they moved in curved paths. In fact, the primary goal of mathematicians, it appears, once the calculus was invented, was to determine curves which solved problems, not just points.

Just to give a quick example, recall that Newton had purportedly shown in the *Principia* that an inverse-square force law implied a conic-section orbit with a focus at the center of force. There was, in fact, great debate in the mathematical community as to whether Newton's sketch was indeed a proof. In any case, Jakob Hermann, among other, used the algebra of differentials to prove this result. Namely, he showed that the inverse-square force law could be translated into the differential equation $-ad^2x = \frac{x(y\,dx - x\,dy)^2}{(x^2+y^2)^{3/2}}$. Hermann then showed, using both algebraic manipulation and Leibniz's techniques for dealing with differentials, that the solution to this equation was given by the equation $a \pm \frac{cx}{b} = \sqrt{x^2 + y^2}$. This was, in fact, the equation of a conic section. As the eighteenth century wore on, algebra grew more and more able to represent paths of motion – and finding such paths became more and more a central problem.

Meanwhile, another question started to arise. How do we know that all the algebraic manipulations we are making are correct? Even in the nineteenth century, there were mathematicians worrying about whether negative numbers made sense, and certainly there was concern about the status of complex numbers. What gradually dawned on practitioners was, that as long as you had some form of axiom system in place, then you could be assured that your calculations gave correct results. So axioms were formulated for arithmetic that were then applied to algebraic manipulations. But when Hamilton discovered the quaternions, mathematicians realized that there could be other sets of axioms that gave interesting results. Perhaps you did not need commutativity of multiplication.

Late in the eighteenth century, another development which would have a major effect in changing the notion of algebra took place. Lagrange, in a major study, tried to determine why Cardano's algebraic solutions of polynomials of degrees three and four could not easily be extended to solutions for polynomials of higher degree. Although Lagrange did not come to any conclusive results, he did introduce the idea of permutations into the search for solutions. His hint was followed up by Abel in the early nineteenth century, who proved that there could be no general algebraic solution of a fifth degree equation. And then Galois developed methods involving what we now call group theory for determining under what conditions polynomial equations are solvable. In fact, Gauss had earlier discovered how to solve a large class of such polynomial equations, those given by cyclotomic polynomials.

Throughout the nineteenth century, the nascent idea of groups grew, particularly once mathematicians realized that many kinds of situations had common properties, including the notion of geometric transformations. In 1854, Cayley gave an axiomatic definition of a group, but his definition did not make it into the mathematical mainstream for another 30 years, when it was essentially rediscovered by Walter Dyck and Heinrich Weber. During the 1890s this definition entered textbooks along with the axiomatic definition of a field, an idea which also had roots in the work of Galois. By the beginning of the twentieth century, algebra had become less about finding solutions to equations and more about looking for common structures in many diverse mathematical objects, with the object being defined by sets of axioms.

This is as far as we will take the history – but to conclude I want to deal somewhat with the pedagogical implications of the history of algebra? Can one use these two sets of stages in the development of the subject in decisions on how to teach algebra in the twenty-first century? Obviously, one cannot answer this question without some research. But I think we should consider, at least, some of the following:

- Should one begin the study of algebraic reasoning by using geometric figures? These are more concrete objects than the x's one usually uses. Squares could be exactly that. Products of numbers can be represented as rectangles. The distributive law is simply a statement about two different ways to represent a given rectangle. And so forth. At the same time, it is probably useful to discuss these concepts verbally. Students are familiar with the "words" of geometry; perhaps they can learn to argue with these words as they draw pictures.
- Should one in the first "real" algebra course (at whatever age this is done) put all one's energy into teaching the solving of problems via equations (of degree one, two, and (perhaps) three)? Right now, I think that most secondary algebra texts are too diffuse. The student does not really know the aim of the course. If one limits the aim to the solution of equations, one has a focus to the course. All algebraic manipulations would then be introduced as tools to solve equations. And when and how to introduce negative numbers should be carefully considered. In addition, one should stress the idea of translating "real-world" problems into algebra.
- Given that the notion of "function" is more abstract than the solution of equations, should we not wait on this until a second course in algebra? The students need lots of experience with curves, which can perhaps be gained in geometry, before dealing with the abstract idea of a function. They also need lots of experience in algebraic manipulation.

- Before one teaches "abstract algebra", namely, groups, rings, fields, and so on, via axiomatic definitions, it is critical that students have enough examples at their fingertips to understand why it is so useful to generalize. They need to understand why particular sets of axioms were chosen. The only way to gain this understanding is through enough experience with examples. It would therefore be reasonable that axiomatics, along with increased abstraction, should appear rather later in the curriculum than earlier.

COMMENTARY FROM A MATHEMATICS EDUCATOR

Reviewing our history, as always, sheds light on contemporary issues. Thus an historical view of the development of algebra has important things to tell us about recent "Numeracy" developments, as well as raising more fundamental questions about appropriate algebra curricula.

The emphasis in many countries on numeracy in primary mathematics education, and its extension into secondary and adult education, makes an assumption that mathematics is to be seen as based on number, and, as such, that algebra is primarily generalized arithmetic. An historical view places number and geometry on at least equal footing in mathematical development, and highlights the powerful interrelationship between the two. Furthermore, we are reminded of the important algebraic role played by symbolism, and of the complexity of the key mathematical ideas of generalization and abstraction.

To the extent that the ontogenetic argument is useful in mathematics education, this reflection should cause us to rethink some current trends. Historically, the motivation for algebra came from the need to solve particular problems, both real-world problems, and those arising from mathematical investigations. Algebra did not arise from an abstract need to generalize arithmetic. The form of algebra today owes more to the nature of its generating problems, and the tools that were used to solve them, than it does to the rules of arithmetic. Educationally, as Katz suggests, it therefore makes sense to consider whether a problem-solving basis would be useful in the early stages of algebraic education. The focus it provides may help to overcome the barrier algebra seems to present for many people.

While algebra has grown, through the later stages of history, into a powerful tool for describing and using mathematical systems, most people will never need to interact with it in this role. Thus setting up algebra as essentially mathematical generalization during compulsory schooling is unlikely to be widely useful, and has not, to date, been pedagogically successful. It usually degenerates into a sequence of algebraic skills that are unrelated to each other and can only be used in recognisable mathematical situations. This is not to ignore the need for school algebra to be able to be turned, at the senior levels, into the study and use of mathematical and axiomatic systems. But, as history teaches us, a problem-solving introduction is consistent with such a development.

I believe that this barrier is also a result of students being unprepared for the type of abstraction that is required to move through generalized arithmetic to manipulating mathematical systems, as current curricula demand. Regarding preparedness for abstraction, again history has a lesson: it takes a long time, and a lot of experience with abstraction, to be comfortable with the particular formal abstractions that characterize modern mathematics. The bringing forward of the grade level at which algebraic ideas are introduced has steadily reduced the opportunities for such practice. Furthermore, it has meant that teachers are working on formal pre-algebra concepts at earlier stages, and using up time that would otherwise be spent on more diverse experiences.

The concentration on structure (as in the New Math initiatives) and on number (as in the Numeracy initiatives) has added to the narrowing of mathematical objects and processes that are part of young children's experiences. Geometric concepts are also vital, and their links with algebra are tangible in a way that numerical concepts are not. To the extent that learning theories are correct about the need for concrete experiences, then the generalizations inherent in drawing and reading maps, building structures, experiencing thrown and falling objects, categorizing characteristics of things, folding paper, and sketching and painting are important components of mathematics education.

Similarly, experiencing multiple uses for numbers and mathematical objects, and multiple ways of representing and describing them, as happened throughout the long history of numbers is essential if sufficient practice at generalization and abstraction are to be achieved. Moving prematurely into the way numbers are dealt with in today's modern mathematics is likely to contribute to difficulties later on.

A final lesson for mathematics education embedded in Katz' description of the development of algebra is the way that concepts we now accept as well-defined had a history of less than exact connotation before taking on their present form. Let us take 'function' as an example. The following is edited from the website: Earliest Uses of Mathematical Words (http://members.aol.com/jeff570/mathword.html, downloaded 15.07.05).

The word **FUNCTION** first appears in a Latin manuscript "Methodus tangentium inversa, seu de fuctionibus" written by Gottfried Wilhelm Leibniz (1646-1716) in 1673. Leibniz used the word in the non-analytical sense, as a magnitude which performs a special duty. He considered a function in terms of "mathematical job"– the "employee" being just a curve. He apparently conceived of a line doing "something" in a given *figura*. . . . From the beginning of his manuscript, however, Leibniz demonstrated that he already possessed the idea of function, a term he denominates *relatio.*

[In] 1692, Leibniz, uses *functiones* in a sense to denote the various 'offices' which a straight line may fulfil in relation to a curve, viz. its tangent, normal, etc.

In ... 1694, ... Leibniz used the word *function* almost in its technical sense, defining *function* as "a part of a straight line which is cut off by straight lines drawn solely by means of a fixed point, and of a point in the curve which is given together with its degree of curvature." The examples given were the ordinate, abscissa, tangent, normal, etc. . . .

[In] 1698, Johann Bernoulli, in another letter to Leibniz, for the first time deliberately assigned a specialized use of the term *function* in the analytical sense. At the end of that month, Leibniz replied (p. 526), showing his approval.

Function is found in English in 1779 in *Chambers' Cyclopedia*: "The term *function* is used in algebra, for an analytical expression any way compounded of a variable quantity, and of numbers, or constant quantities"

Thus, not only was function developed through geometry, but also it took much use before its analytical sense was fully developed. It is an open, and interesting, question whether an introduction to the idea of function as a tool, possibly a geometric tool as Katz suggests, would be a useful curricula approach.

There is probably no one "best" way to teach algebra. But it is clear that since elementary algebra is the key to any success in mathematics at all and abstract algebra is critical to work in advanced mathematics, we must increase the flow of students through our algebra courses. I believe that attention to the history of the field is a valuable guide in deciding how to do this. I hope that these ideas can be taken seriously in the discussions of the future of the teaching and learning of algebra.

NOTE: Further historical details on the events mentioned in this paper can be found in either (Katz, 1998) or (Katz, 2004).

References

Al-Khwarizmi, M.: 1831, *The Algebra of Mohammed ben Musa*, Frederic Rosen, trans., Oriental Translation Fund, London.

Cardano, G.: 1968, *The Great Art, or The Rules of Algebra*, T. Richard Witmer, trans., MIT Press, Cambridge.

Euclid: 1993, *The Data of Euclid*, George McDowell and Merle Sokolik, trans., Union Square Press, Baltimore.

Euclid: 2002, *Euclid's Elements*, The Thomas Heath translation, edited by Dana Densmore, Green Lion Press, Santa Fe.

Euler, L.: 1984, *Elements of Algebra*, translated by John Hewlett (reprinted from the 1840 edition), Springer-Verlag, New York.

Galilei, G.: 1974, *Two New Sciences*, Stillman Drake, trans., University of Wisconsin Press, Madison.

Katz, V.: 1998, *A History of Mathematics: An Introduction*, 2nd ed., Addison-Wesley, Reading, MA.

Katz, V.: 2004, *A History of Mathematics, Brief Edition*, Addison-Wesley, Boston.

Kepler, J.: 1992, *New Astronomy*, William H. Donahue, trans., Cambridge University Press, Cambridge.

Maclaurin, C.: 1748, *A Treatise of Algebra in Three Parts*, A. Millar, London.

Neugebauer, O. and Sachs, A.: 1945, *Mathematical Cuneiform Texts*, American Oriental Society, New Haven.

Puig, L.: 2004, 'The history of algebra in mathematics education,' in Kaye Stacey, Helen Chick, and Margaret Kendal, eds., *The Future of the Teaching and Learning of Algebra*, Kluwer, Boston, pp. 189–223.

Recorde, R.: 1969, *The Whetstone of Witte*, photographic reprint edition, Da Capo Press, New York.

VICTOR J. KATZ
University of the District of Columbia Washington
DC, USA
841 Bromley St. Silver Spring
MD 20902, 301-649-5729
E-mail: vkatz@udc.edu

Commentary from a Mathematics Educator

BILL BARTON
The University of Auckland
Auckland, New Zealand
E-mail: b.barton@auckland.ac.nz