

Chapter 1

Quadratic Equations

1.1 Introduction

Since the solution of a linear equation $aX = b$ does not use anything more than a division, it hardly belongs to the algebraic theory of equations; it is therefore appropriate to begin these lectures with quadratic equations

$$aX^2 + bX + c = 0 \quad (a \neq 0).$$

Dividing both sides by a , we may reduce this to

$$X^2 + pX + q = 0.$$

The solution of this equation is well-known: when $\left(\frac{p}{2}\right)^2$ is added to each side, the square of $X + \frac{p}{2}$ appears and the equation can be written

$$\left(X + \frac{p}{2}\right)^2 + q = \left(\frac{p}{2}\right)^2.$$

(This procedure is called “completion of the square”). The values of X easily follow:

$$X = -\frac{p}{2} \pm \sqrt{\left(\frac{p}{2}\right)^2 - q}.$$

This formula is so well-known that it may be rather surprising to note that the solution of quadratic equations could not have been written in this form before the seventeenth century.* Nevertheless, mathematicians had been solving quadratic

*The first uniform solution for quadratic equations (regardless of the signs of coefficients) is due to Simon Stevin in “L’Arithmetique” [55, p. 595], published in 1585. However, Stevin does not use literal coefficients, which were introduced some years later by François Viète: see chapter 4, §4.1.

equations for about 40 centuries before. The purpose of this first chapter is to give a brief outline of this “prehistory” of the theory of quadratic equations.

1.2 Babylonian algebra

The first known solution of a quadratic equation dates from about 2000 B.C.; on a Babylonian tablet, one reads (see Van der Waerden [62, p. 69])

I have subtracted from the area the side of my square: 14.30.
Take 1, the coefficient. Divide 1 into two parts: 30. Multiply 30
and 30: 15. You add to 14.30, and 14.30.15 has the root 29.30.
You add to 29.30 the 30 which you have multiplied by itself: 30,
and this is the side of the square.

This text obviously provides a procedure for finding the side of a square (say x) when the difference between the area and the side (i.e. $x^2 - x$) is given; in other words, it gives the solution of $x^2 - x = b$.

However, one may be puzzled by the strange arithmetic used by Babylonians. It can be explained by the fact that their base for numeration is 60; therefore 14.30 really means $14 \cdot 60 + 30$, i.e. 870. Moreover, they had no symbol to indicate the absence of a number or to indicate that certain numbers are intended as fractions. For instance, when 1 is divided by 2, the result which is indicated as 30 really means $30 \cdot 60^{-1}$, i.e. 0.5. The square of this 30 is then 15 which means 0.25, and this explains why the sum of 14.30 and 15 is written as 14.30.15: in modern notations, the operation is $870 + 0.25 = 870.25$.

After clearing the notational ambiguities, it appears that the author correctly solves the equation $x^2 - x = 870$, and gets $x = 30$. The other solution $x = -29$ is neglected, since the Babylonians had no negative numbers.

This lack of negative numbers prompted Babylonians to consider various types of quadratic equations, depending on the signs of coefficients. There are three types in all:

$$X^2 + aX = b, \quad X^2 - aX = b, \quad \text{and} \quad X^2 + b = aX,$$

where a, b stand for positive numbers. (The fourth type $X^2 + aX + b = 0$ obviously has no (positive) solution.)

Babylonians could not have written these various types in this form, since they did not use letters in place of numbers, but from the example above and from other numerical examples contained on the same tablet, it clearly appears that the

Babylonians knew the solution of

$$X^2 + aX = b \quad \text{as} \quad X = \sqrt{\left(\frac{a}{2}\right)^2 + b} - \frac{a}{2}$$

and of

$$X^2 - aX = b \quad \text{as} \quad X = \sqrt{\left(\frac{a}{2}\right)^2 + b} + \frac{a}{2}.$$

How they argued to get these solutions is not known, since in every extant example, only the procedure to find the solution is described, as in the example above. It is very likely that they had previously found the solution of geometric problems, such as to find the length and the breadth of a rectangle, when the excess of the length on the breadth and the area are given. Letting x and y respectively denote the length and the breadth of the rectangle, this problem amounts to solving the system

$$\begin{cases} x - y = a \\ xy = b. \end{cases} \quad (1.1)$$

By elimination of y , this system yields the following equation for x :

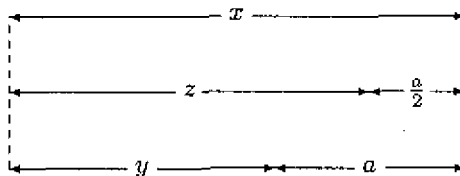
$$x^2 - ax = b. \quad (1.2)$$

If x is eliminated instead of y , we get

$$y^2 + ay = b. \quad (1.3)$$

Conversely, equations (1.2) and (1.3) are equivalent to system (1.1) after setting $y = x - a$ or $x = y + a$.

They probably deduced their solution for quadratic equations (1.2) and (1.3) from their solution of the corresponding system (1.1), which could be obtained as follows: let z be the arithmetic mean of x and y .



In other words, z is the side of the square which has the same perimeter as the given rectangle:

$$z = x - \frac{a}{2} = y + \frac{a}{2}.$$

Compare then the area of the square (i.e. z^2) to the area of the rectangle ($xy = b$). We have

$$xy = \left(z + \frac{a}{2}\right)\left(z - \frac{a}{2}\right)$$

whence $b = z^2 - \left(\frac{a}{2}\right)^2$. Therefore, $z = \sqrt{\left(\frac{a}{2}\right)^2 + b}$ and it follows that

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

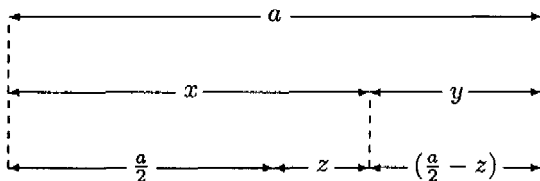
This solves at once the quadratic equations $x^2 - ax = b$ and $y^2 + ay = b$.

Looking at the various examples of quadratic equations solved by Babylonians, one notices a curious fact: the third type $x^2 + b = ax$ does not explicitly appear. This is even more puzzling in view of the frequent occurrence in Babylonian tablets of problems such as to find the length and the breadth of a rectangle when the perimeter and the area of the rectangle are given; which amounts to the solution of

$$\begin{cases} x + y = a \\ xy = b. \end{cases} \quad (1.4)$$

By elimination of y , this system leads to $x^2 + b = ax$. So, why did Babylonians solve the system (1.4) and never consider equations like $x^2 + b = ax$?

A clue can be discovered in their solution of system (1.4), which is probably obtained by comparing the rectangle with sides x, y to the square with perimeter $\frac{a}{2}$:



One then sets $x = \frac{a}{2} + z$, whence $y = \frac{a}{2} - z$, and finishes as before.

Whatever their method, the solution they get is

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

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

thus assigning one value for x and one value for y , while it is clear to us that x and y are interchangeable in the system (1.4): we would have given two values for each one of the unknown quantities, and found

$$x = \frac{a}{2} \pm \sqrt{\left(\frac{a}{2}\right)^2 - b}, \quad y = \frac{a}{2} \mp \sqrt{\left(\frac{a}{2}\right)^2 - b}.$$

In the Babylonian phrasing, however, x and y are not interchangeable: they are the length and the breadth of a rectangle, so there is an implicit condition that $x \geq y$. According to S. Gandz [22, §9], the type $X^2 + b = aX$ was systematically and purposely avoided by Babylonians because, unlike the two other types, it has *two* positive solutions (which are the length x and the breadth y of the rectangle). The idea of two values for one quantity was probably very embarrassing to them, it would have struck Babylonians as an illogical absurdity, as sheer nonsense.

However, this observation that algebraic equations of degree higher than 1 have several interchangeable solutions is of fundamental importance: it is the corner-stone of Galois theory, and we shall have the opportunity to see to what clever use it will be put by Lagrange and later mathematicians.

As André Weil commented in relation to another topic [69, p. 104]

This is very characteristic in the history of mathematics. When there is something that is really puzzling and cannot be understood, it usually deserves the closest attention because some time or other some big theory will emerge from it.

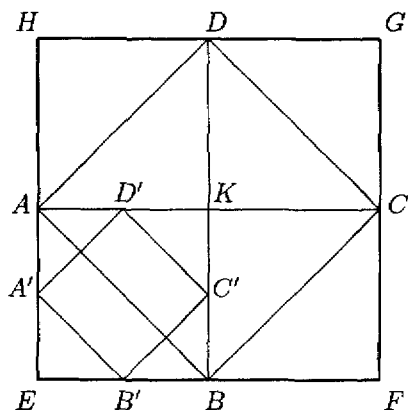
1.3 Greek algebra

The Greeks deserve a prominent place in the history of mathematics, for being the first to perceive the usefulness of proofs. Before them, mathematics were rather empirical. Using deductive reasoning, they built a huge mathematical monument, which is remarkably illustrated by Euclid's celebrated masterwork "The Elements" (c. 300 B.C.).

The Greeks' major contribution to algebra during this classical period is foundational. They discovered that the naive idea of number (i.e. integer or rational number) is not sufficient to account for geometric magnitudes. For instance, there is no line segment which could be used as a length unit to measure the diagonal and the side of a square by integers: the ratio of the diagonal to the side (i.e. $\sqrt{2}$) is not a rational number, or in other words, the diagonal and the side are incommensurable.

The discovery of irrational numbers was made among followers of Pythagoras, probably between 430 and 410 B.C. (see Knorr [39, p. 49]). It is often credited to Hippasus of Metapontum, who was reportedly drowned at sea for producing a downright counterexample to the Pythagoreans' doctrine that "all things are numbers." However, no direct account is extant, and how the discovery was made is still a matter of conjecture. It is widely believed that the first magnitudes which were shown to be incommensurable are the diagonal and the side of a square, and the following reconstruction of the proof has been proposed by Knorr [39, p. 27]:

Assume the side AB and the diagonal AC of the square $ABCD$ are both measured by a common segment; then AB and AC both represent numbers (= integers) and the squares on them, which are $ABCD$ and $EFGH$, represent square numbers. From the figure, it is clear (by counting triangles) that $EFGH$ is the double of $ABCD$, so $EFGH$ is an even square number and its side EF is therefore even. It follows that EB also represents a number, whence $EBKA$ is a square number.



Since the square $ABCD$ clearly is the double of the square $EBKA$, the same arguments show that AB is even, whence $A'B'$ represents a number.

We now see that $A'B'$ and $A'C'$ ($= EB$), which are the halves of AB and AC , both represent numbers; but $A'B'$ and $A'C'$ are the side and the diagonal of a new (smaller) square, so we may repeat the same arguments as above.

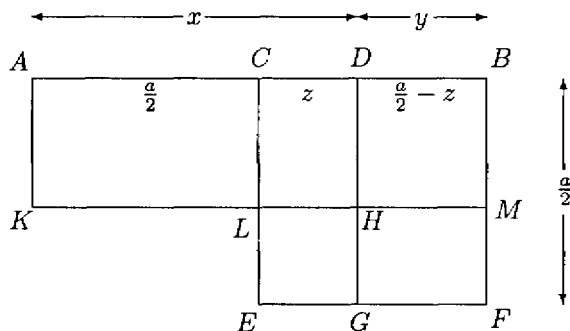
Iterating this process, we see that the numbers represented by AB and AC are indefinitely divisible by 2. This is obviously impossible, and this contradiction proves that AB and AC are incommensurable.

This result obviously shows that integers are not sufficient to measure lengths of segments. The right level of generality is that of ratios of lengths. Prompted by this discovery, the Greeks developed new techniques to operate with ratios of geometric magnitudes in a logically coherent way, avoiding the problem of assigning numerical values to these magnitudes. They thus created a "geometric algebra," which is methodically taught by Euclid in "The Elements."

By contrast, Babylonians seem not to have been aware of the theoretical difficulties arising from irrational numbers, although these numbers were of course unavoidable in the treatment of geometric problems: they simply replaced them by rational approximations. For instance, the following approximation of $\sqrt{2}$ has been found on some Babylonian tablet: $1.24.51.10$, i.e. $1 + 24 \cdot 60^{-1} + 51 \cdot 60^{-2} + 10 \cdot 60^{-3}$ or $1.41421296296296 \dots$, which is accurate up to the fifth place.

Although Euclid does not explicitly deal with quadratic equations, the solution of these equations can be detected under a geometric garb in some propositions of the Elements. For instance, Proposition 5 of Book II states [30, v. I, p. 382]:

If a straight line be 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.



On the figure above, the straight line AB has been cut into equal segments at C and unequal segments at D , and the proposition asserts that the rectangle AH together with the square LG (which is equal to the square on CD) is equal to the square CF . (This is clear from the figure, since the rectangle AL is equal to the rectangle DF).

If we understand that the unequal segments in which the given straight line $AB = a$ is cut are unknown, it appears that this proposition provides us with the core of the solution of the system

$$\begin{cases} x + y = a \\ xy = b. \end{cases}$$

Indeed, setting $z = x - \frac{a}{2}$ "the straight line between the points of section," it states that $b + z^2 = \left(\frac{a}{2}\right)^2$. It then readily follows that

$$z = \sqrt{\left(\frac{a}{2}\right)^2 - b}$$

whence

$$x = \frac{a}{2} + \sqrt{\left(\frac{a}{2}\right)^2 - b} \quad \text{and} \quad y = \frac{a}{2} - \sqrt{\left(\frac{a}{2}\right)^2 - b},$$

as in Babylonian algebra. In subsequent propositions, Euclid also teaches the solution of

$$\begin{cases} x - y = a \\ xy = b \end{cases}$$

which amounts to $x^2 - ax = b$ or $y^2 + ay = b$. He returns to the same type of problems, but in a more elaborate form, in propositions 28 and 29 of book VI (Compare Kline [38, pp. 76–77] and Van der Waerden [62, p. 121].)

The Greek mathematicians of the classical period thus reached a very high level of generality in the solution of quadratic equations, since they considered equations with (positive) *real* coefficients. However, geometric algebra, which was the only rigorous method of operating with real numbers before the XIXth century, is very difficult. It imposes tight limitations which are not natural from the point of view of algebra; for instance, a great skill in the handling of proportions is required to go beyond degree three.

To progress in the theory of equations, it was necessary to think more about formalism and less about the nature of coefficients. Although later Greek mathematicians such as Hero and Diophantus took some steps in that direction, the

really new advances were brought by other civilizations. Hindus, and Arabs later, developed techniques of calculation with irrational numbers, which they treated unconcernedly, without worrying about their irrationality. For instance, they were familiar with formulas like

$$\sqrt{a} + \sqrt{b} = \sqrt{a + b + 2\sqrt{ab}}$$

which they obtained from $(u + v)^2 = u^2 + v^2 + 2uv$ by extracting roots of both sides and replacing u and v by \sqrt{a} and \sqrt{b} respectively. Their notion of mathematical rigor was rather more relaxed than that of Greek mathematicians, but they paved the way to a more formal (or indeed algebraic) approach to quadratic equations (see Kline [38, ch. 9, §2]).

1.4 Arabic algebra

The next landmark in the theory of equations is the book “Al-jabr w’ al muqabala” (c. 830 A.D.), due to Mohammed ibn Musa al-Khowarizmi.

The title refers to two basic operations on equations. The first is *al-jabr* (from which the word “algebra” is derived) which means “the restoration” or “making whole.” In this context, it stands for the restoration of equality in an equation by adding to one side a negative term which is removed from the other. For instance, the equation

$$x^2 = 40x - 4x^2$$

is converted into

$$5x^2 = 40x$$

by *al-jabr* [36, p. 105]. The second basic operation *al muqabala* means “the opposition” or “balancing”; it is a simplification procedure by which like terms are removed from both sides of an equation. For instance, *al muqabala* changes

$$50 + x^2 = 29 + 10x$$

into

$$21 + x^2 = 10x$$

[36, p. 109].

In this work, al-Khowarizmi initiates what might be called the classical period in the theory of equations, by reducing the old methods for solving equations to a few standardized procedures. For instance, in problems involving several unknowns, the systematically sets up an equation for one of the unknowns, and he solves the three types of quadratic equations

$$X^2 + aX = b, \quad X^2 + b = aX, \quad X^2 = aX + b$$

by completion of the square, giving the two (positive) solutions for the type $X^2 + b = aX$.

Al-Khowarizmi first explains the procedure, as a Babylonian would have done:

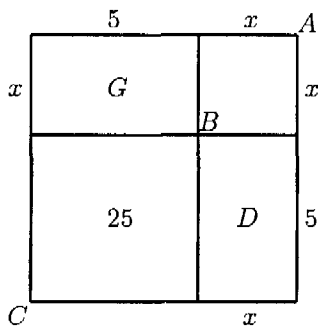
The following is an example of squares and roots equal to numbers: a square and 10 roots are equal to 39 units. The question therefore in this type of equation is about as follows: what is the square which combined with ten of its roots will give a sum total of 39? The manner of solving this type of equation is to take one-half of the roots just mentioned. Now the roots in the problem before us are 10. Therefore take 5, which multiplied by itself gives 25, an amount which you add to 39, giving 64.

Having taken then the square root of this which is 8, subtract from it the half of the roots, 5, leaving 3. The number three therefore represents one root of this square, which itself, of course, is 9. Nine therefore gives that square. [36, pp. 71–73]

However, after explaining the procedure for solving each of the six types $mX^2 = aX$, $mX^2 = b$, $aX = b$, $mX^2 + aX = b$, $mX^2 + b = aX$ and $mX^2 = aX + b$, he adds:

We have said enough, says Al-Khowarizmi, so far as numbers are concerned, about the six types of equations. Now, however, it is necessary that we should demonstrate geometrically the truth of the same problems which we have explained in numbers. [36, p. 77]

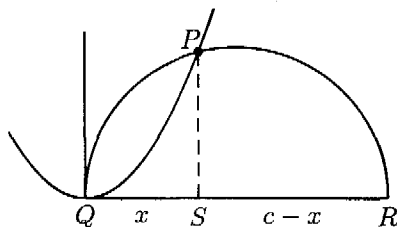
He then gives geometric justifications for his rules for the last three types, using completion of the square as in the following example for $x^2 + 10x = 39$:



Let x^2 be the square AB . Then $10x$ is divided into two rectangles G and D , each being $5x$ and being applied to the side x of the square AB . By hypothesis, the value of the shape thus produced is $x^2 + 10x = 39$. There remains an empty corner of value $5^2 = 25$ to complete the square AC . Therefore, if 25 is added, the square $(x + 5)^2$ is completed, and its value is $39 + 25 = 64$. It then follows that $(x + 5)^2 = 64$, whence $x + 5 = 8$ and $x = 3$ (see [36, p. 81]).

It should be observed that the geometry behind this construction is much more elementary than in Euclid's *Elements*, since it is not logically connected by deductive reasoning to a small number of axioms, but relies instead on intuitive geometric evidence. From the point of view of algebra, on the other hand, al-Khwarizmi's work is incommensurately ahead of Euclid's, and it set the stage for the later development of algebra as an independent discipline.

Another remarkable achievement of the Arabs in the theory of equations is a geometric solution of cubic equations due to Omar Khayyam (c. 1079). For instance, the solution of $x^3 + b^2x = b^2c$ is obtained by intersecting the parabola $x^2 = by$ with the circle of diameter c which is tangent to the axis of the parabola at its vertex.



To prove that the segment x as shown on the figure above satisfies $x^3 + b^2x = b^2c$,

we start from the relation $x^2 = b \cdot PS$, which yields

$$\frac{b}{x} = \frac{x}{PS}.$$

On the other hand, since the triangles QSP and PSR are similar, we have

$$\frac{x}{PS} = \frac{PS}{c - x},$$

whence

$$\frac{b}{x} = \frac{PS}{c - x}.$$

As $PS = b^{-1}x^2$, this equation yields

$$\frac{b}{x} = \frac{x^2}{b(c - x)},$$

whence $x^3 = b^2c - b^2x$, as required.

Omar Khayyam also gives geometric solutions for the other types of cubic equations by intersection of conics, but these brilliant solutions are of little use for practical purposes, and an algebraic solution was still longed for.

In 1494, Luca Pacioli closes his book "Summa de Arithmetica, Geometria, Proportione e Proportionalita" (one of the first printed books in mathematics) with the remark that the solutions of $x^3 + mx = n$ and $x^3 + n = mx$ (in modern notations) are as impossible as the quadrature of the circle. (See Kline [38, p. 237], Cardano [11, p. 8].) However, unexpected developments were soon to take place.