CHAPTER 5

Algebra: The Search for an Elusive Formula

5.1 Introduction

On the night of February 12, 1535, Niccolò Tartaglia, of Brescia, found the solution to the following vexing problem: "A man sells a sapphire for 500 ducats, making a profit of the cube root of his capital. How much is his profit?" This triumph helped him gain victory over Antonio Maria Fiore, who had posed this problem to challengers as part of a public contest. Besides fame, the prize for the winner included thirty banquets prepared by the loser for Tartaglia and his friends. (Tartaglia chose to decline this part of the prize.) Such contests were part of academic life in sixteenth-century Italy, as competitors vied for university positions and sponsorship from the nobility [93, p. 329].

In order to find a solution to the above problem Tartaglia needed to solve the equation

$$x^3 + x = 500,$$

where x is an unknown real number. The method to solve cubic equations of this type had been discovered by Scipione del Ferro, professor at the University of Bologna, who had passed on the secret to his pupil Fiore. Del Ferro's solution for cubic equations of the form $x^3 + cx = d$, in modern notation, amounts to the formula

$$x = \sqrt[3]{\sqrt{(d/2)^2 + (c/3)^3} + d/2} - \sqrt[3]{\sqrt{(d/2)^2 + (c/3)^3} - d/2}.$$

Before continuing, the reader is invited to verify that this formula, with $c = 1, d = 500$, does indeed produce an exact solution to the contest equation, which had given so much trouble to Tartaglia. (Note that Tartaglia did not

PHOTO 5.1. Babylonian problem text on tablet YBC 4652.

have a calculator.) A further list of problems that Fiore posed to Tartaglia can be found in [58, p. 254].

Much earlier, around 2000 B.C.E., the Babylonians had solved problems that can be interpreted as quadratic equations [91, pp. 108 ff.]. In fact, many Babylonian clay tablets have been preserved with lists of mathematical problems on them. Scholars have generally assumed that the motivation for trying to solve quadratic equations originally arose from the need to solve practical or scientific problems. Alternative interpretations are suggested in [93, p. 34].

Following is a typical example of such a problem, taken from [173, p. 63] (see also [128]). The numbers, in the sexagesimal (base 60) system the Babylonians used, are given in the notation of [128]. The number $63\frac{1}{2}$, for instance, is represented as 1,3;30, where the semicolon plays the role of the decimal point. Likewise, $1;3,30 = 1\frac{7}{120}$. See [119, pp. 162 ff.] for a discussion of the Babylonian number system. We follow [173, pp. 63 ff.] in the interpretation of the tablet that is numbered AO 8862, dating from the Hammurabi dynasty, ca. 1700 B.C.E.

Length, width. I have multiplied length and width, thus obtaining the area. Then I added to the area, the excess of the length over the width: $3,3$ (i.e. 183 was the result). Moreover, I have added length and width: 27. Required length, width and area.

(given:) 27 and $3,3$, the sums

(result:) 15 length $\quad$ 3, 0 area.
12 width

One follows this method:

$$27 + 3, 3 = 3, 30$$
$$2 + 27 = 29.$$

Take one half of 29 (this gives 14; 30).

$$14; 30 \times 14; 30 = 3, 30; 15$$
$$3, 30; 15 - 3, 30 = 0; 15.$$

The square root of 0; 15 is 0; 30.

$$14; 30 + 0; 30 = 15 \quad \text{length}$$
$$14; 30 - 0; 30 = 14 \quad \text{width}.$$

Subtract 2, which has been added to 27, from 14, the width. 12 is the actual width. I have multiplied 15 length by 12 width.

$$15 \times 12 = 3, 0 \quad \text{area}$$
$$15 - 12 = 3$$
$$3, 0 + 3 = 3, 3.$$

Translated into modern notation, the problem amounts to solving a system of two equations

$$xy + x - y = 183,$$
$$x + y = 27.$$

The author makes a change of variable to transform the system into the standard form in which such problems were solved. If we set

$$y' = y + 2,$$

then the system gets transformed into

$$xy' = 210,$$
$$x + y' = 29.$$

The solution of a general system of the form

$$xy' = b,$$
$$x + y' = a,$$

follows a standard recipe, given by the three equations

$$w = \sqrt{(a/2)^2 - b},$$
$$x = (a/2) + w,$$
$$y' = (a/2) - w.$$

Rather than giving the general solution method in this way, it was indicated by a series of examples, all solved in this fashion.

Observe that solving our system $xy' = b, x + y' = a$ is equivalent to solving the equation $(a - x)x = ax - x^2 = b$. The Babylonians thus had a method to solve certain quadratic equations. Even some cubic and quartic equations, as well as some systems with as many as ten unknowns, were within their reach. (See also Exercises 5.1 and 5.2.) The solutions were phrased in essentially numerical terms, without recourse to geometry (in contrast to later Greek mathematics), even though quantities were represented as line segments, rectangles, etc. Also, the solutions were given in the form of a procedure, or algorithm, without any indication as to why the procedure provided the correct solution.

Beginning with the Pythagoreans (approx. 500 B.C.E.), a class of problems known as "application of areas" appeared in Greek mathematics. At around the same time, Chinese mathematics was developing elaborate methods for solving simultaneous systems of linear equations [91, pp. 173 f.][116, pp. 124, 249 ff.], and Indian Vedic mathematics developed accurate methods of calculating roots, and considered, but did not solve, certain quadratic equations [91, p. 273].

The most important type of Greek application of areas problem was the following: Given a line segment, construct on part of it, or on the line segment extended, a parallelogram equal in area to a given rectilinear figure and falling short (in the first case) or exceeding (in the second case) by a parallelogram similar to a given parallelogram [97, p. 34]. One may view the systems of equations solved by the Babylonians as application of areas problems. Indeed, suppose we want to solve the system

$$\begin{aligned} x + y &= a, \\ xy &= b^2. \end{aligned}$$

Using the terminology of Figure 5.1, we wish to apply to $AB(= a)$ a rectangle $AH(= ax - x^2)$ equal to the given area b^2, and falling short of $AM(= ax)$ by the square BH. This is equivalent to solving the quadratic equation

$$x(a - x) = b^2.$$

The first original source we consider in this chapter is Proposition 5 from Book II of Euclid's *Elements*, written around 300 B.C.E. Together with the Pythagorean Theorem, it gives a method to solve such quadratic equations, that is, to construct the solutions with straightedge and compass. In the general statement of application of areas problems above, the "falling short" case was termed *ellipsis* and the "exceeding" case *hyperbolē*. Here lies the origin of the terminology used for the conic sections [97, pp. 91–92].

FIGURE 5.1. Application of areas.

Euclid does not explicitly treat quadratic equations. Book II is a rather short part of the *Elements* and contains a collection of geometric propositions, which Zeuthen [181] calls “geometric algebra.” A typical example is the following (Proposition 1): *If there be two straight lines, and one of them be cut into any number of segments whatever, the rectangle contained by the two straight lines is equal to the rectangles contained by the uncut straight line and each of the segments.* Translated into algebraic notation, this corresponds to the formula

$$a(b + c + \cdots) = ab + ac + \cdots .$$

In essence, Book II might be viewed as the beginning of a textbook on algebra, written in geometrical language, and as a continuation of Babylonian algebra [173, p. 119]. Later, the Arab scholars Ibn Qurra (908–946) and Abu Kamil (ca. 850–930) used results from Book II to solve equations, illustrating the correspondence between algebra and geometry [12, Ch. 4].

Ironically, one reason for the lack of progress on higher-degree equations was the very rigor that so distinguishes classical Greek mathematics. The restriction of rigorous mathematics to geometry forced the use of geometric methods with their inherent complications and limitations. It also made it impossible to consider equations of degree higher than three in the absence of a geometric interpretation. Major progress in the theory of equations was made during the “Silver Age” of classical Greek mathematics (about 250–350 C.E.) by Diophantus of Alexandria. He systematically introduced symbolic abbreviations for the terms in an equation, while the Egyptians and Babylonians used prose. In addition, he was the first to consider exponents higher than the third. For more details see [93, Sec. 5.2].

Another reason for the lack of further progress in solving algebraic equations of degree higher than two was complications involved in the development of a rigorous treatment of irrational numbers as mathematical objects, exemplified by Books V and X of the *Elements*. It was known that not all geometric magnitudes are rational, such as the diagonal in a square of side length one (Exercise 5.3). To avoid dealing with the nature of such irrational magnitudes Euclid employs geometric algebra with elaborate proportionality arguments [97, pp. 173 ff.], [169, pp. 15–16]. Although solutions to certain cubic equations arise in the work of Archimedes (287–212 B.C.E.) and Menaechmus (middle of the fourth century B.C.E.) (see [93, pp. 108–109]), it was left to later cultures to make significant progress toward methods for solving higher-degree equations.

A first important step was accomplished by mathematics in India, during the period 200–1200 C.E. Less rigorous than Greek mathematics and much less tied to geometry, it allowed a correct arithmetic of negative and irrational numbers. This approach prevented the obsession with philosophical difficulties that prevented much progress in its Greek predecessor [97, pp. 184 ff.]. (For more details see Chapter 6 of [93].)

The rise of Islam during the seventh century and its spread from the Arabian peninsula all the way to countries bordering China in one direction and to Spain in the other direction prepared the ground for important developments in our story. In the eighth century, the newly built city of Baghdad, the present-day capital of Iraq, became one of the political and scientific centers of the Islamic world. In a setting described in the tales of *The Thousand and One Nights*, the Caliph Harun al-Rashid had a large library constructed, containing many scientific works written in Sanskrit, Persian, and Greek. During the ninth century, the Caliph al-Ma'mun supplemented this library with a new intellectual center called the House of Wisdom. Devoted to research and a massive translation project, it led to an influx of scholars from all parts of the Islamic sphere of influence, as well as a growing collection of original manuscripts together with translations into Arabic. By the end of the century, the major works of Euclid, Archimedes, Apollonius, Diophantus, and other Greek mathematicians were available to Islamic scholars, in addition to many manuscripts of Babylonian and Indian origin. (See [12] for a detailed description of Islamic mathematics.)

One of the scholars in the House of Wisdom was al-Khwarizmi (ca. 800–847), who around 825 wrote *The Condensed Book on the Calculation of al-Jabr and al-Muqābala.* He explains the purpose of the treatise in the introduction:

> That fondness for science, by which God has distinguished the Imam al-Ma'mun, the Commander of the Faithful... has encouraged me to compose a short work on calculating by *al-jabr* and *al-muqābala*, confining it to what is easiest and most useful in arithmetic, such as men constantly require in cases of inheritance, legacies, partition, law-suits, and trade, and in all their dealings with one another, or where measuring of lands, the digging of canals, geometrical computation, and other objects of various sorts and kinds are concerned [93, p. 229].

From the title of this work derives our word *algebra* to denote the branch of mathematics that Europe later learned from it. The following interpretation of the words in the title is given in [93, p. 228].

> The term *al-jabr* can be translated as "restoring" and refers to the operation of transposing a subtracted quantity on one side of an equation to the other side where it becomes an added quantity. The word *al-muqābala* can be translated as "comparing" and refers to the reduction of a positive term by subtracting equal amounts from both sides of the equation. For example, converting $3x + 2 = 4 - 2x$ to $5x + 2 = 4$ is an example of *al-jabr* while converting the latter to $5x = 2$ is an example of *al-muqābala*.

The text contains a classification of different types of quadratic equations, with only positive coefficients. Subsequently, a collection of numerical recipes is given to solve the different types. Influenced by Greek tradition, al-Khwarizmi then gives geometric proofs for his procedures.

In studying the Greek geometry texts, one encounters problems that lead to cubic equations. The problem of "doubling the cube" was one such instance. This was the problem of doubling the cubical altar at Delos, reportedly posed by an oracle, for which the Delians sought Plato's help [93, p. 42]. The correct interpretation, according to Plato (427–348/347 B.C.E.), was to construct with straightedge and compass a cube that has twice the volume of the altar cube. If we assume the side length of the altar cube to be 1, then we are led to the problem of constructing the cube root of 2, which was shown to be impossible only in the nineteenth century by Pierre Wantzel (1814–1848) [93, p. 598]. An excellent reference for more details on these problems is [98].

If in addition to straightedge and compass one allows the use of conic sections, such as parabolas and hyperbolas, however, then problems such as this one have ready solutions. See [93, pp. 108–9]. See also [12, Ch. 3] for Islamic contributions to constructibility problems. Several Islamic mathematicians applied conic sections to other cubic equations as well. The most systematic effort in this direction is due to Omar Khayyam (1048–1131), who classified all types of cubic equations and provided solutions for them, using conic sections. In his work *Treatise on Demonstrations of Problems of al-Jabr and al-Muqabala*, he emphasizes that the reader needs to be thoroughly familiar with the work of Euclid, Apollonius, and al-Khwarizmi in order to follow the solutions. For instance, a solution to the equation $x^3 + cx = d$ is found by intersecting a circle and a parabola [93, p. 244]. An English translation of part of the *Treatise* can be found in [122].

In Persia, Khayyam's treatise was used as a school text for centuries. In the Western world Khayyam became known for his poetry through the 1859 publication of an English translation of *Rubaiyat of Omar Khayyam*. This book went through more than 300 editions and spawned a plethora of secondary works. The year 1900 saw the foundation of the Omar Khayyam Club of America in Boston, and there is continued interest in his poetry [122, pp. 583–584].

Despite Khayyam's ingenuity, his work did not lead to progress in the theory of equations, since his methods relied heavily on geometry. The world still had to wait for an algebraic method applicable to higher-degree equations. The time was ripe for such a method in sixteenth century Renaissance Italy. A number of mathematicians were working on algebraic methods for solving cubic equations. Still lacking a workable concept of negative numbers, the usual categorization was used depending on which side of the equation contained which terms (with nonnegative coefficients). As mentioned at the beginning of this section, it was Scipione del Ferro (1465–1526), a professor at the University of Bologna, who first found an algebraic method to deal with the equation $x^3+cx = d$ (Exercises 5.4–5.5). Niccolò Tartaglia (1499–1557) claimed to possess a solution for the equation $x^3 + bx^2 = d$. This caught the attention of a Milanese mathematician, Girolamo Cardano (1501–1576). Cardano was working on an arithmetic text, and invited Tartaglia to tell him about his methods, to be included in the book, with full credit to Tartaglia. Finally, according to Tartaglia himself, Tartaglia was willing to tell Cardano what his "rule" was for finding the solutions of his cubic, but not his "method," which would explain why the rule indeed produced the right solutions. In modern terms, he was willing to give Cardano the algorithm to compute the solutions, without any indication why the algorithm indeed produced the desired result, nor how one could discover the algorithm. Cardano swore not to divulge Tartaglia's rule. However, he began working on the problem himself afterwards, and published his solution in his greatest work, the *Ars Magna* (The Great Art) [29]. There followed a long public dispute with Tartaglia over Cardano's supposed breach of confidence, told in some detail in [93, Ch. 9]. The second text we will study in this chapter is the solution of one of the cases of the cubic out of the *Ars Magna*, the equation $x^3 = bx + d$, or, in Cardano's words, the case of "the cube equal to the first power and number."

While most of the *Ars Magna* is devoted to solving the cubic, it also contains a method for solving equations of degree 4. This method had been discovered by Ludovico Ferrari (1522–1565), a student of Cardano. Cardano included Ferrari's solution in his book but relegated it to the last of its forty chapters. As he says about it in the introduction:

> Although a long series of rules might be added and a long discourse given about them, we conclude our detailed consideration with the cubic, others being merely mentioned, even if generally, in passing. For as *positio* [the first power] refers to a line, *quadratum* [the square] to a surface, and *cubum* [the cube] to a solid body, it would be very foolish for us to go beyond this point. Nature does not permit it [29, p. 9].

Ferrari had proceeded by first making a change of variable in the general fourth-degree equation in order to eliminate the cubic term. If our equation is

$$x^4 + ax^3 + bx^2 + cx + d = 0,$$

then the substitution

$$y = x + \frac{a}{4}$$

results in the equation

$$y^4 + py^2 + qy + r = 0,$$

for the appropriate p, q, r. Now rewrite this equation as

$$\left(y^2 + \frac{p}{2}\right)^2 = -qy - r + \left(\frac{p}{2}\right)^2.$$

Adding a quantity u to the equation inside the parentheses on the left-hand side, we obtain

$$\left(y^2 + \frac{p}{2} + u\right)^2 = -qy - r + \left(\frac{p}{2}\right)^2 + 2uy^2 + pu + u^2.$$

Now, Ferrari's clever idea was to determine a constant u, depending on p and q, but not on y, that would make the right-hand side a perfect square of a first degree polynomial in y. This leads to the cubic equation

$$8u^3 + 8pu^2 + (2p^2 - 8r)u - q^2 = 0,$$

which can now be solved by Cardano's method (Exercise 5.6). This equation is now called a *resolvent cubic* for the quartic we started with, and we will have occasion to return to it in the work of Lagrange in the eighteenth century.

The formulas for the roots of third- and fourth-degree equations raised as many questions as they answered, the hallmark of any important mathematical discovery. Just how many roots did an equation have? Why did Cardano's formula produce some solutions and not others? For instance, the equation $x^3 + 16 = 12x$ has the solution $x = 2$, but Cardano's formula produces another solution, $x = -4$. And, most importantly, why was it sometimes necessary to extract square roots of negative numbers to evaluate the formula, even when the end result was a real number? Finally, one was forced to take negative and imaginary numbers seriously as mathematical objects. Before, they had been dismissed as absurd and unnatural. It was not too long before the algebra of complex numbers appeared in the influential text *Algebra* by Rafael Bombelli (1526–1573), published in 1572.

Other areas of mathematics, such as analytic geometry and the differential calculus, began to take center stage at the beginning of the seventeenth century, and further progress in the theory of algebraic equations was slow in coming. Cardano's work pushed the rudimentary system of algebraic notation to its utmost limit, and real progress required a more user-friendly symbolism, which was slow in developing. First steps in this direction were taken by François Viète (1540–1603). (See also the number theory chapter.)

After such great success with cubic and quartic equations, the next obvious step was to attack the general quintic equation. Here progress was to be slow, however. While several new methods were developed for cubic and quartic equations, none of them resulted in a method for the quintic. The next milestone that we will look at does not appear until the second half of the eighteenth century, in the work of Joseph Louis Lagrange (1736–1813).[1] Lagrange's 1771 memoir *Réflexions sur la Résolution Algébrique*

[1] Lagrange's contributions to mathematics are woven into the strands of two other chapters in this book as well. He taught an analysis course at the Ecole Polytechnique, through which he discovered that one of his pupils who submitted

des Equations (Reflections on the Solution of Algebraic Equations) [101] is one of the gems to be found in this subject and was crucial for the eventual solution of the problem. Algebra was synonymous with the theory of equations, and Lagrange's work is the outstanding contribution to the subject during the second half of the eighteenth century. In a leisurely style, the eminently readable and lengthy work first gives an analysis of all the work done on the problem, surveys the different methods of Cardano, Euler, etc., and attempts to extract from them general principles:

> These methods all come down to the same general principle, knowing how to find functions of the roots of the given equation, which are such 1° that the equation or equations by which they are given, i.e., of which they are roots (equations which are commonly called the reduced equations), themselves have degree less than that of the given equation, or at least are factorable into other equations of degree less that that; 2° that one can easily compute from these the values of the desired roots [94, p. 45].

He makes it clear that any such method will involve in an essential way the study of certain types of permutations of the roots of the equation, a conviction borne out by the later course of events. (A permutation of the roots of an equation can be thought of simply as a shuffle, or rearrangement, of the roots.) The memoir contains a number of results about such permutations, which foreshadow the later emergence of group theory.[2] After many very insightful observations, he concludes with a rather pessimistic assessment of the possibility for finding a general algebraic method for solving equations of all degrees. Here we will study extensive excerpts from Lagrange's memoir to illustrate these points.

The next major advance had to wait until the emergence of new generations of mathematicians toward the beginning of the nineteenth century. First there was the Italian Paolo Ruffini (1765–1822), who in 1799 published a lengthy treatise *General Theory of Equations* [146, v. 1, pp. 1–324]. It contained a proof that the general equation of degree five was not solvable by radicals, meaning that there could be no formula for the roots that involved only algebraic operations on the coefficients of the equation together with radicals of these coefficients. Due to the less than clear exposition, his

---

particularly brilliant homework assignments had them in fact done by Sophie Germain, who later made important contributions to Fermat's Last Theorem. See the number theory chapter.

[2] A group is a set, together with a binary operation that combines two group elements to give a third. This operation is required to be associative, have an identity element, and each element has an inverse. An example is the set of integers with addition as the binary operation. Group theory is the study of such structures with the ultimate goal of creating a classification scheme.

PHOTO 5.2. Abel.

arguments were received with much skepticism (later on a significant gap was found), and were never accepted by the mathematical community [169, pp. 273 ff.].

In 1824, the Norwegian mathematician Niels Henrik Abel (1802–1829) gave a different proof, published in the first issue of *Journal für die Reine und Angewandte Mathematik* in 1826 [1, vol. 1, pp. 66–94]. There is some evidence [67] that Abel had similar results for higher-degree equations of prime degree, as well as results on the roots of equations that are solvable by radicals. In addition, he achieved a characterization of a large class of equations that in fact do admit a solution by radicals. This class includes all so-called cyclotomic equations, discussed below [169, p. 302]. Possibly only his untimely death at the age of 26 prevented him from being the one to give a complete solution to the problem of which algebraic equations are solvable by radicals. This honor was reserved for the Frenchman Evariste Galois (1811–1832), one of the most flamboyant characters in the history of mathematics.

At the tender age of 18, Galois communicated to the Academy of Sciences in Paris some of his results on the theory of equations, through one of its members, Augustin-Louis Cauchy (1789–1857). Shortly thereafter, Galois learned that a number of his results had actually been obtained by Abel, before him. Two years later, he submitted a rewritten version, *Mémoire sur les Conditions de Résolubilité des Equations par Radicaux* (Memoir on the Conditions for Solvability of Equations by Radicals), which, as it turned out much later, contained the punch line of our story. With Abel's work it had become clear that the search for a general method for solving algebraic equations by radicals was doomed. There remained the problem of characterizing those equations that *are* amenable to such a solution. The last source in this chapter consists of excerpts from Galois's memoir. They show his work to be a new kind of mathematics that would be characteristic of the nineteenth century, and which marked a gradual departure from the work of Lagrange, to which it nevertheless owed its inspiration. Galois's characterization of algebraic equations solvable by radicals forms one of the beginnings of abstract algebra, namely the theory of groups and that of fields. Most of a standard first-semester graduate course on abstract algebra is nowdays devoted to the legacy of Galois's groundbreaking work. The importance of his memoir was not recognized properly until 1843, when Joseph Liouville prepared Galois's manuscripts for publication and announced that Galois had indeed solved this age-old problem.

Even though much has been omitted in this summary account, it cannot end without mentioning the work of Carl Friedrich Gauss (1777–1855) on cyclotomic equations, published in 1801 in his great treatise *Disquisitiones Arithmeticae* (Arithmetical Investigations). Gauss's work was the culmination of another strand of our story, which will have to receive short shrift, but needs to be outlined in any account of the theory of equations.

As is often the case in mathematics, developments in other fields were ultimately to play a big role in the theory of equations as well. In 1702, the German mathematician and philosopher Gottfried Wilhelm Leibniz (1646–1716) published a paper in the scientific journal *Acta Eruditorum* with the curious title *New Specimen of the Analysis for the Science of the Infinite about Sums and Quadratures* [110]. In this paper, Leibniz discusses what we now call the method of partial fractions, in particular the possibility of decomposing a rational function, that is, a quotient of two polynomial functions, into a sum of fractions with denominators that are either linear or quadratic polynomials. This would reduce the integration of rational functions to integrating only those functions involving $1/x$ and $1/(x^2+1)$, which Leibniz calls the "quadrature of the hyperbola and the circle"[3] (see

[3] In order to understand why Leibniz relates $1/(x^2+1)$ to the quadrature (area) of the unit circle, consider the following figure:

Now, using calculus, one can show that the area of the circle segment enclosed by the angle θ is given by

$$\frac{\theta}{2} = \frac{\tan^{-1}(x)}{2} = \frac{1}{2}\int_0^x \frac{dt}{t^2+1}.$$

PHOTO 5.3. *Disquisitiones Arithmeticae.*

the analysis chapter). To obtain such decompositions, he embarks on an investigation of the factorization of real polynomials into irreducible factors, hoping to show that every polynomial with real coefficients can be factored into irreducible factors of degree one or two. This is in fact true, and is now known as the "Fundamental Theorem of Algebra." Leibniz, however, ends up making an embarrassing mistake and comes up with the following "counterexample." From

$$x^4 + a^4 = \left(x^2 + a^2\sqrt{-1}\right)\left(x^2 - a^2\sqrt{-1}\right),$$

it follows that

$$x^4+a^4=\left(x + a\sqrt{-\sqrt{-1}}\right)\left(x - a\sqrt{-\sqrt{-1}}\right)\left(x + a\sqrt{\sqrt{-1}}\right)\left(x - a\sqrt{\sqrt{-1}}\right),$$

from which he erroneously concludes that no nontrivial combination of these linear factors can result in a real divisor of $x^4 + a^4$. Moreover, he thought that expressions like $\sqrt{\sqrt{-1}}$ led to a whole new type of number:

> Therefore, $\int dx/(x^4 + a^4)$ cannot be reduced to the squaring of the circle or the hyperbola by our analysis above, but founds a new kind of its own [110, p. 360] [169, p. 99].

Of course, there was no need, in this case, to get into any investigations about new numbers at all. As Nikolaus Bernoulli (1687–1759) pointed out a little later, one easily solves this problem by observing that

$$x^4 + a^4 = (x^2 + a^2 + \sqrt{2}ax)(x^2 + a^2 - \sqrt{2}ax).$$

Apparently, Isaac Newton (1642–1727) had obtained the same factorization some time earlier. The mystery of extracting roots of complex numbers also was resolved eventually. In 1739, the French mathematician Abraham de Moivre (1667–1754) showed that radicals of complex numbers do not produce a new kind of number. Given an "impossible binomial $a + \sqrt{-b}$," in his terminology, he showed that if φ is an angle such that $\cos(\varphi) = a/\sqrt{a^2 + b}$, then the nth roots of $a + \sqrt{-b}$ are

$$\sqrt[2n]{a^2 + b}\left(\cos(\psi) + \sqrt{\cos^2(\psi) - 1}\right),$$

where ψ ranges over the n values $\varphi/n, (2\pi - \varphi)/n, (2\pi + \varphi)/n, (4\pi - \varphi)/n, (4\pi + \varphi)/n, \ldots$. But this expression is again of the form $a + \sqrt{-b}$ for some other values of a and b. Thus, Leibniz's objection was refuted, and subsequently serious attempts were undertaken to prove the Fundamental Theorem of Algebra. A number of proofs were given, all containing more or less serious flaws. As with several other important results, it was left to Gauss to correct all errors and fill all gaps to give the first essentially complete proof in 1799. Subsequently, he gave several different proofs of the result. An equivalent and more common formulation of it is that any non-constant polynomial with complex coefficients factors into linear factors. (See [169, Ch. 7] for a more detailed discussion.)

Another accomplishment of Gauss needs to be mentioned, which was of great significance in validating some of Lagrange's conclusions. De Moivre's work on roots of complex numbers showed that the nth root of a number may be far from unique. A special role is assigned to the nth roots of 1, since one can easily see that if r is an nth root of a complex number s, and $\omega_1, \ldots, \omega_{n-1}$ are the distinct nth roots of 1 other than 1 itself, then the other nth roots of s are of the form $\omega_1 r, \ldots, \omega_{n-1} r$. Once again it was left to Gauss, in a brilliant piece of work published as part of the *Disquisitiones Arithmeticae*, to tell the whole story about these roots of unity, as they are commonly called. In this he built on previous work by a number of other mathematicians, notably Vandermonde, Euler, and Lagrange. The nth roots of unity are of course all roots of the polynomial $x^n - 1$, which we can factor into

$$x^n - 1 = (x - 1)(x^{n-1} + x^{n-2} + \cdots + x + 1).$$

It was shown earlier that it was sufficient to study the matter when n is prime. Gauss showed that if n is prime, then the second factor is irreducible over the rational numbers and has $n - 1$ distinct complex roots. Most importantly, these *cyclotomic equations*, as they are called, are all solvable by radicals (Exercise 5.7). Thus, Gauss had found the first instance of an important class of equations of arbitrarily high degree that are solvable by radicals. Shortly afterwards, Abel and Galois undertook the Herculean task of classifying precisely those equations that, like the cyclotomic ones, are amenable to this method of solution.

As happens often in mathematics, in the end the tools developed in order to attack the problem of solving equations by radicals took on far greater importance than the actual solution itself. We already mentioned the consequences of Galois's work for the development of abstract algebra, in particular the theory of groups and the theory of field extensions, which have become central to all of algebra as well as to applications of algebra to other subjects inside and outside of mathematics. In the case of group theory, a thorough account can be found in [179], while [48] and [174] contain excellent accounts of the historical development of Galois theory. Most importantly, Galois was advocating a paradigm shift, away from a purely computational approach, in favor of an abstract qualitative analysis of the problems of algebra. As he says in an unpublished preface to his memoir:

> Now, I think that the simplifications produced by the elegance of calculations (intellectual simplifications, I mean; there are no material simplifications) are limited; I think the moment will come where the algebraic transformations foreseen by the speculations of analysts will not find nor the time nor the place to occur any more; so that one will have to be content with having foreseen them. . . .
>
> Jump above calculations; group the operations, classify them according to their complexities rather than their appearances; this, I believe, is the mission of future mathematicians; this is the road on which I am embarking in this work [66, p. 9] [169, p. 397].

The next such paradigm shift was initiated by Emmy Noether (1882–1935) in the 1920s, and gave the subject of algebra the form it has today. She introduced a new level of abstraction into algebra, which led to great conceptual clarity and formal elegance, but also removed the subject somewhat from its origins. As a consequence, hardly anyone talks about equations and solution by radicals anymore when discussing Galois theory, which has become a beautiful

PHOTO 5.4. Noether.

axiomatic theory concerned with extensions of fields and the theory of groups.

Lest the reader become too complacent in marveling at the good fortune of such a complete and satisfying solution to the important and longstanding problem of solving polynomial equations, a rather rare occurrence in mathematics, there is of course still an unanswered question lingering, which opens up entirely new vistas in seemingly distant parts of mathematics. Given that the roots of an algebraic equation of degree five or greater cannot in general be expressed as algebraic functions of the coefficients of the equation, one might wonder whether there are other kinds of functions that give a representation of the roots, and if so, which ones. The fruitfulness of this question is born out by the book *Vorlesungen über das Ikosaeder* (Lectures on the Icosahedron), published in 1884 by the great nineteenth-century German mathematician Felix Klein (1849–1925) [95], in which he investigates these questions for equations of degree five. As the title suggests, unexpected connections to geometry appear, and his work has inspired much present-day research on this and related problems.

Exercise 5.1: Use the Babylonian method described in this section to solve the system

$$\begin{aligned} 2xy + x - 4y &= 1 \\ x + 4y &= 2. \end{aligned}$$

What is the general principle behind the substitution to transform a given system into standard form?

Exercise 5.2: The Babylonians also solved systems of linear equations with a method similar to the one discussed in this section. Solve the following problem (see [128, p. 66]): From one field I have harvested 4 gur of grain per bùr (surface unit). From a second field I have harvested 3 gur of grain per bùr. The yield of the first field was 8,20 more than that of the second. The areas of the two fields were together 30,0 bùr. How large were the fields?

Exercise 5.3: Use the Pythagorean Theorem and the Fundamental Theorem of Arithmetic to show that the diagonal in a square of side length one cannot be a rational number.

Exercise 5.4: Use del Ferro's method to solve the following problem from Fiore's list of challenge problems that he gave to Tartaglia: "There is a tree, 12 braccia high, which was broken into two parts at such a point that the height of the part which was left standing was the cube root of the length of the part that was cut away. What was the height of the part that was left standing?"

Exercise 5.5: Use del Ferro's formula for the roots of a cubic to solve the equation $x^3 = 15x + 4$. (Note that this equation has three real roots.)

Exercise 5.6: Work out the details in Ferrari's method of solving an equation of degree four. For hints check [169, Sect. 3.2].

Exercise 5.7: Find all roots of the polynomial $x^4 - 1$.

5.2 Euclid's Application of Areas and Quadratic Equations

As mentioned in the introduction, Book II of Euclid's *Elements* is rather short, with only fourteen propositions. It contains a collection of results from "geometric algebra." And indeed a number of the results translate into well-known formulas in modern algebra. The original source we discuss in this section is Proposition 5. It turns out to be the essential ingredient in the solution of certain types of quadratic equations, aside from its use in