# Solving the Cubic Equation and Beyond

Frank Sottile

We discuss formulas for the solutions to polynomial equations in one variable, such as

$$3x + 1 = 2\,,\ 2x^2 + 1 = 4x\,,\ x^3 = 7x + 6\,, \dots\,.$$

Letters near the end of the alphabet $s, t, x, y, \dots$ will be our variables, but we will also use letters such as $a, b, c, A, B, C, \alpha, \beta, \gamma, \dots$ to represent real number constants, including the coefficients. We seek formulas for the solutions in terms of the coefficients.

**Beginnings.** What is the simplest type of equation to solve?

Of course the answer is a linear equation, one of the form $ax + b = 0$, whose solution is $x = -b/a$. Solutions to such equations are discussed in the Rhind Papyrus (c.1650 BCE) and probably are significantly more ancient.

How about the next simplest type of equation?

That would be a quadratic equation, one of the form $ax^2 + bx + c = 0$, whose solution is given by the familiar quadratic formula

$$x \;=\; \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}\,.$$

We derive this by completing a square. To begin, divide $ax^2 + bx + c = 0$ by its leading coefficient to get $x^2 + bx/a + c/a = 0$, and then set $\beta := \frac{b}{a}$ and $\gamma = -\frac{c}{a}$ to get $x^2 + \beta x = \gamma$. Let us try to simplify this using $x = y + \delta$, which gives

$$(y + \delta)^2 + \beta(y + \delta) = \gamma \quad \text{or} \quad y^2 + (2\delta + \beta)y + \delta^2 + \beta\delta = \gamma\,.$$

A propitious choice is $\delta = -\beta/2$, for that removes the linear term, and we get

$$y^2 \;=\; \gamma - \delta^2 - \beta\delta \;=\; \gamma + \frac{\beta^2}{4}\,. \tag{1}$$

This leads to $y = \pm\sqrt{\gamma + \beta^2/4}$. Substituting back for $x$, and then for $a$, $b$, and $c$ gives the familiar formula

$$x \;=\; -\frac{\beta}{2} \pm \sqrt{\gamma + \frac{\beta^2}{4}} \;=\; -\frac{b}{2a} \pm \sqrt{\frac{b^2}{4a^2} - \frac{c}{a}}\,.$$

Before moving on, I ask you: why is there a $\pm$ ? (Answer $1^2 = (-1)^2 = 1$, so 1 has two square roots. Using both square roots of 1 gives the two solutions to (1).) This method of *completing the square* goes back to the Babylonians in 2000–1600 BCE, and increasingly more sophisticated methods were employed by Euclid (geometric), and a more algebraic method was given by Brahmagupta (628 CE). This modern form goes back to René Descartes (1596–1650 CE).

**Solution to the cubic.** The next most complicated equation is a cubic equation. Scipione del Ferro (1465–1526) discovered the solution (which is for a special case) we will give, but told no one. At that time, one got and kept one's academic job by winning public challenges, demonstrating an ability to solve problems that others could not[1]. On his death bed, he passed the formula to his son-in-law Annibale della Nave (c. 1500–1558). Earlier, he had shared this secret with his student Antonio Maria Fior, who later challenged Niccolò Fontana (1499–1577) (a.k.a. Tartaglia, the stammerer) to a contest, posing to Tartaglia the

[1] Think of a peacock's display to signaling its fitness.

solutions of many cubic equations. Under tremendous pressure, Tartaglia discovered the formula himself (on 12 February 1535) and vanquished Fior, cementing his reputation.

Later (1539), Gerolamo Cardano (1501–1576) talked the formula out of Tartaglia, but swore an oath of secrecy. With his student Ludovico Ferrari (1522–1565), they discovered the general formula. They undertook serious library research on a trip to Bologna in 1543, with the breakthrough coming when Annibale della Nave shared with them his father-in-laws notes. Then Cardano published this and much more in his book, *Ars Magna*, in 1545. This is one of the greatest scientific treatises of the early Renaissance.

Let us discuss solving the cubic equation $Ax^3 + Bx^2 + Cx + D = 0$. For now, consider solving an easier one, say

$$Ax^3 + D = 0 \;\rightsquigarrow\; x^3 = -D/A =: d \;\rightsquigarrow\; x = \sqrt[3]{d}\,.$$

What is $\sqrt[3]{d}$? In Calculus, we show that the function $f(x) = x^3$ is a bijection; every real number is a cube of a real number in a unique way, and so the cube root of any real number is a well defined real number.

While that seems settled, we can ask if this is the only solution to $x^3 = d$. Let us explore this when $d = 1$. Consider the following

$$\begin{aligned}(-1+\sqrt{-3})^3 \;&=\; (-1)^3 + 3(-1)^2\sqrt{-3} + 3(-1)(\sqrt{-3})^2 + (\sqrt{-3})^3\\ &=\; -1 + 3\sqrt{-3} + 9 - 3\sqrt{-3} \;\;=\;\; 8 \;=\; 2^3\,.\end{aligned}$$

Thus $1 = (-\frac{1}{2} + \frac{\sqrt{-3}}{2})^3$ and a similar calculation (or taking complex conjugates), gives $1 = (-\frac{1}{2} - \frac{\sqrt{-3}}{2})^3$. Let us define

$$\zeta \;:=\; -\tfrac{1}{2} \;+\; \tfrac{\sqrt{-3}}{2}\,. \tag{2}$$

Then $\zeta^2 = -\frac{1}{2} - \frac{\sqrt{-3}}{2}$, so that $\zeta^3 = 1 = (\zeta^2)^3$. These, together with 1, are the three *cube roots of 1* or *cube roots of unity*.

In this way there are three solutions to the equation $x^3 = d$ for $d \in \mathbb{R}$, namely

$$x \;=\; \sqrt[3]{d}\,,\; \zeta\sqrt[3]{d}\,,\; \text{and } \zeta^2\sqrt[3]{d}\,.$$

We return to the general cubic equation $Ax^3 + Bx^2 + CX + D = 0$. First, let us divide by $A$ to simplify and get $x^3 + bx^2 + cx + d = 0$, where $b = B/A$, $c = C/A$, and $d = D/A$. This time the substitution $x = y + \delta$ is most propitious for $\delta = -b/3$ for that will cancel the coefficient $b$ of $x^2$ (you should check this). Substituting $x = y - b/3$, gives

$$y^3 + y(c - \tfrac{b^2}{3}) + d - \tfrac{bc}{3} + \tfrac{2b^3}{27} \;=\; 0\,.$$

Let us rewrite (and thus simplify) this as

$$y^3 + \beta y \;=\; \gamma\,, \tag{3}$$

which is called the *depressed cubic*, a reduction due to Cardano and Ferrari.

The first step towards solving this is involves an amazing idea. Note that

$$(s+t)^3 \;=\; s^3 + 3s^2t + 3st^2 + t^3 \;=\; 3st(s+t) + (s^3 + t^3)\,,$$

or

$$(s+t)^3 - 3st(s+t) \;=\; s^3 + t^3\,.$$

Comparing this to (3), pattern matching suggests the interesting substitutions

$$y = s + t\,,\ \beta = -3st\,,\ \text{and } \gamma = s^3 + t^3\,.$$

Solving for $t$ in $\beta = -3st$ and substituting this in the formula for $\gamma$ gives

$$t \;=\; -\frac{\beta}{3s} \qquad \rightsquigarrow \qquad \gamma \;=\; s^3 - \frac{\beta^3}{27s^3}\,.$$

Clearing denominators gives

$$s^6 \;-\; \gamma s^3 \;-\; \frac{\beta^3}{27} \;=\; 0\,.$$

While this may seem a step backwards (reducing a cubic to a sextic!), it is in fact a quadratic in disguise, a quadratic in $s^3$. Solve using the quadratic formula to get

$$s^3 \;=\; \frac{\gamma}{2} \pm \sqrt{\frac{\gamma^2}{4} + \frac{\beta^3}{27}}\,. \tag{4}$$

One solution to the sextic is then

$$s \;=\; \sqrt[3]{\frac{\gamma}{2} + \sqrt{\frac{\gamma^2}{4} + \frac{\beta^3}{27}}}\,. \tag{5}$$

Writing $t^3 = \gamma - s^3$ gives $t^3 = \frac{\gamma}{2} - \sqrt{\frac{\gamma^2}{4} + \frac{\beta^3}{27}}$. Using $y = s + t$ gives,

$$y \;=\; \sqrt[3]{\frac{\gamma}{2} + \sqrt{\frac{\gamma^2}{4} + \frac{\beta^3}{27}}} \;+\; \sqrt[3]{\frac{\gamma}{2} - \sqrt{\frac{\gamma^2}{4} + \frac{\beta^3}{27}}}\,.$$

Note that we use both solutions to (4) here.

**Example.** Let us try to solve $y^3 + 6y = 20$ using this method. Here, $\beta = 6$ and $\gamma = 20$, and the innermost quantity is $(\frac{\gamma}{2})^2 + (\frac{\beta}{3})^3 = 10^2 + 2^3 = 108 = 3 \cdot 6^2$, and so we obtain

$$y \;=\; \sqrt[3]{10 + 6\sqrt{3}} \;+\; \sqrt[3]{10 - 6\sqrt{3}}\,.$$

This may be simplified, as the terms $10 \pm 6\sqrt{3}$ are perfect cubes. For that, try the Ansatz[2]:

$$\begin{aligned}(a + b\sqrt{3})^3 \;&=\; a^3 + 3a^2b\sqrt{3} + 3a(b\sqrt{3})^2 + (b\sqrt{3})^3 \\ &=\; a^3 + 3a^2b\sqrt{3} + 9ab^2 + 3b^3\sqrt{3} \;\;=\;\; a(a^2 + 9b^2) \;+\; 3b(b^2 + a^2)\sqrt{3}\,.\end{aligned}$$

So we are looking for solutions in $a, b$ to

$$10 \;=\; a(a^2 + 9b^2) \qquad \text{and} \qquad 6 = 3b(b^2 + a^2)\,.$$

We may find a solution $a = b = 1$ by inspection, and so we obtain that

$$y \;=\; 1 + \sqrt{3} \;+\; 1 - \sqrt{3} \;= 2\,,$$

which we check is a solution to our equation as $2^3 + 6 \cdot 2 = 20$.

What about the other roots? Long division of $y^3 + 6y - 20$ by $y - 2$ gives the factorization

$$y^3 + 6y - 20 \;=\; (y - 2)(y^2 + 2y + 10)\,,$$

[2]An Ansatz is an educated guess that is verified later by its results

and then the quadratic formula gives the other two roots as

$$-\frac{2}{2} \pm \sqrt{\frac{2^2}{4} - 10} \;=\; -1 \pm \sqrt{-9} \;=\; -1 \pm 3\sqrt{-1}\,. \tag{6}$$

Let us check this,

$$(-1 \pm 3\sqrt{-1})^3 + 6(-1 \pm 3\sqrt{-1}) \;=\; -1 \pm 9\sqrt{-1} + 27 \mp 27\sqrt{-1} \;-6 \pm 18\sqrt{-1} \;=\; 20\,.$$

There is one other way to obtain the remaining roots, it is a bit of a sleight-of-hand. Recall the equation $\beta = -3st$, and the cube root of unity (2), $\zeta$. Inserting $1 = \zeta^3$ between $s$ and $t$, we have $\beta = -3s\zeta^3 t = -3(\zeta s)(\zeta^2 t) = -3(\zeta^2 s)(\zeta t)$. Note then that $(\zeta s)^3 = (\zeta^2 s)^3 = s^3$, and the same for $t$, so that we still have

$$\gamma \;=\; s^3 + t^3 \;=\; (\zeta s)^3 + (\zeta^2 t)^3 \;=\; (\zeta^2 s)^3 + (\zeta t)^3\,.$$

Recall that (4) gives three solutions for $s$, the one in (5) and the following two:

$$s \;=\; \zeta\sqrt[3]{\frac{\gamma}{2} + \sqrt{\frac{\gamma^2}{4} + \frac{\beta^3}{27}}} \qquad \text{and} \qquad s \;=\; \zeta^2\sqrt[3]{\frac{\gamma}{2} + \sqrt{\frac{\gamma^2}{4} + \frac{\beta^3}{27}}}\,,$$

The corresponding values for $t$ that we get from the expressions for $\beta$ by associating $\zeta^3$ between the factors of the original $s$ and $t$ are

$$t \;=\; \zeta^2\sqrt[3]{\frac{\gamma}{2} - \sqrt{\frac{\gamma^2}{4} + \frac{\beta^3}{27}}} \qquad \text{and} \qquad t \;=\; \zeta\sqrt[3]{\frac{\gamma}{2} - \sqrt{\frac{\gamma^2}{4} + \frac{\beta^3}{27}}}\,.$$

These give two other possibilities for the solution $y$,

$$\begin{aligned} y_2 \;&=\; \zeta\sqrt[3]{\frac{\gamma}{2} + \sqrt{\frac{\gamma^2}{4} + \frac{\beta^3}{27}}} \;+\; \zeta^2\sqrt[3]{\frac{\gamma}{2} - \sqrt{\frac{\gamma^2}{4} + \frac{\beta^3}{27}}}\,, \qquad \text{and} \\ y_3 \;&=\; \zeta^2\sqrt[3]{\frac{\gamma}{2} + \sqrt{\frac{\gamma^2}{4} + \frac{\beta^3}{27}}} \;+\; \zeta\sqrt[3]{\frac{\gamma}{2} - \sqrt{\frac{\gamma^2}{4} + \frac{\beta^3}{27}}}\,. \end{aligned} \tag{7}$$

Let us now return to $y^3 + 6y - 20 = 0$. These formulas (7) give

$$\begin{aligned} y_2 \;&=\; \left(-\tfrac{1}{2} + \tfrac{\sqrt{-3}}{2}\right)(1+\sqrt{3}) \;+\; \left(-\tfrac{1}{2} - \tfrac{\sqrt{-3}}{2}\right)(1-\sqrt{3}) \;=\; -1 + 3\sqrt{-1} \\ y_3 \;&=\; \left(-\tfrac{1}{2} - \tfrac{\sqrt{-3}}{2}\right)(1+\sqrt{3}) \;+\; \left(-\tfrac{1}{2} + \tfrac{\sqrt{-3}}{2}\right)(1-\sqrt{3}) \;=\; -1 - 3\sqrt{-1}\,, \end{aligned}$$

the same roots that we found in (6).

Let us solve another cubic,

$$y^3 - 15y \;=\; 4\,.$$

For this, $\beta/3 = -5$ and $\gamma/2 = 2$, and so

$$y \;=\; \sqrt[3]{2 + \sqrt{4 - 125}} \;+\; \sqrt[3]{2 - \sqrt{4-125}} \;=\; \sqrt[3]{2 + 11\sqrt{-1}} \;+\; \sqrt[3]{2 - 11\sqrt{-1}}\,.$$

To find the cube root, let us try this Ansatz,

$$\begin{aligned} (a + b\sqrt{-1})^3 \;&=\; a^3 + 3a^2b\sqrt{-1} - 3ab^2 - b^3\sqrt{-1} \\ &=\; a(a^2 - 3b^2)^3 + b(3a^2 - b)\sqrt{-1} \quad \stackrel{?}{=} \; 2 + 11\sqrt{-1}\,. \end{aligned}$$

A little playing with numbers shows that $a = 2$ and $b = 1$ works, so that a solution is

$$y \;=\; (2 + \sqrt{-1}) \;+\; (2 - \sqrt{-1}) \;=\; 4\,.$$

Before returning to the human story, let us have a few words about numbers.

**Negative numbers.** While some mathematicians had discussed and even used negative numbers by the 1540's, they were not considered real or legitimate. There was no subtraction sign, and Cardano avoided negative numbers in his book (except as a tool for solving simultaneous linear equations).

**Complex numbers.** These were avoided even more. (This is why we often call the number $\sqrt{-1}$ 'imaginary'.) Cardano talked of the 'mental tortures' involved when he used them (sparingly) in *Ars Magna*.

Despite this, mathematicians eventually came around to the use of complex and imaginary numbers, in part because of the rôle of imaginary numbers in the del Ferro-Tartaglia formula, particularly to find the 'real' solution $y = 4$ for the cubic equation $y^3 = 15y + 4$, as we saw above. In his 1572 book *L'Algebra*, Bombelli introduced $\sqrt{-1}$ and used it to revisit Cardano's work, including the solution of the cubic $y^3 = 15y + 4$.

Back to the human story... Tartaglia was incensed by what he viewed as Cardano's betrayal in *Ars Magna*, where Cardano revealed the del Ferro-Tartaglia formula. He challenged Cardano to a contest. Cardano declined. Then Tartaglia challenged Ferrari to a contest in 1548, which he lost.

Ferrari solved the quartic in 1540. We skip his solution in favor of a more interesting one. We begin with a depressed quartic equation

$$y^4 + \alpha y^2 + \beta y + \gamma \;=\; 0\,. \tag{8}$$

Suppose for the moment that we find numbers $q, r, s, t$ such that

$$y^4 + \alpha y^2 + \beta y + \gamma \;=\; (y^2 + qy + r)(y^2 + sy + t)\,. \tag{9}$$

Then we would be able to find all roots to the quartic (8). (Why would we be done?) Expanding the right hand side of the factorization (9) gives

$$y^4 + (q+s)y^3 + (r + qs + t)y^2 + (rs + qt)y + rt\,.$$

Equating coefficients with the quartic (8) gives the four equations,

$$0 \;=\; q + s\,,\ \alpha \;=\; r + qs + t\,,\ \beta \;=\; rs + qt\,,\ \text{and}\ \gamma \;=\; rt\,.$$

Solving the first for $q$ and substituting gives

$$q \;=\; -s\,,\ \alpha + s^2 \;=\; r + t\,,\ \beta \;=\; s(r - t)\,,\ \text{and}\ \gamma \;=\; rt\,.$$

The following expressions allow us to eliminate $r$ and $t$,

$$\begin{aligned} s^2(\alpha + s^2)^2 - \beta^2 \;&=\; s^2(r+t)^2 - s^2(r-t)^2 \\ &=\; 4s^2rt \;=\; 4s^2\gamma\,. \end{aligned}$$

Notice that this uses $n^2 - m^2 = (n+m)(n-m)$, where $n = r + t$ and $m = r - t$, thus $(r+t)^2 - (r-t)^2 = (r+t+r-t)(r+t-(r-t)) = 4rt$. Using that $(\alpha+s^2)^2 = \alpha^2 + 2\alpha s^2 + s^4$, expanding and collecting, we get the sextic,[3]

$$s^6 \;+\; 2\alpha s^4 \;+\; (\alpha^2 - 4\gamma)s^2 \;-\; \beta^2 \;=\; 0\,. \tag{10}$$

Setting $\sigma := s^2$ gives a cubic in $\sigma$. We use the del Ferro-Tartaglia formula to solve for $\sigma$ and then take the square roots to get $s$ and $q = -s$. Then we may use these solutions in $\beta = s(r - t)$ and $\gamma = rt$ to solve for $r$ and $t$ (these will be a quadratic), and then this will

[3]Again a sextic. What does 6 have to do with 4?

give us our factorization (9). Finally, two applications of the quadratic formula gives us the solutions to the depressed quartic (8) with which we began. Note that each of the three solutions for $\sigma$ to the cubic gives a different factorization with the sign $\pm s$ coming from solving $s^2 = \sigma$ interchanging factors. This can be written out to give a 'formula' for the solution to a quartic, which is somewhat involved. (It takes about a page to write down.) We will leave our solution as above, a series of reductions.

At this point, you might want to try your hand at this formula/algorithm on a real example. Let us try this with the extremely depressed quartic equation $y^4 - 5y - 6 = 0$. Then the cubic associated to (10) is

$$\sigma^3 \ + \ 24\sigma \ - \ 25 \ = \ 0\,. \tag{11}$$

The del Ferro-Tartaglia for this cubic (with $\gamma = 25$ and $\beta = 24$) gives

$$\sigma \ = \ \sqrt[3]{\tfrac{25}{2} + \sqrt{\tfrac{625}{4} + 512}} \ + \ \sqrt[3]{\tfrac{25}{2} - \sqrt{\tfrac{625}{4} + 512}} \ = \ \sqrt[3]{\tfrac{25}{2} + \tfrac{9}{2}\sqrt{33}} \ + \ \sqrt[3]{\tfrac{25}{2} - \tfrac{9}{2}\sqrt{33}}\,.$$

The simplification is because $\frac{625}{4} + 512 = 2673/4$ and $2673 = 81 \cdot 33$. A similar Ansatz as before for $(a + b\sqrt{33})^3 = \frac{25}{2} \pm \frac{9}{2}\sqrt{33}$ gives $a = b = \frac{1}{2}$, and so

$$s^2 \ = \ \sigma \ = \ \tfrac{1}{2} + \tfrac{1}{2}\sqrt{33} \ + \ \tfrac{1}{2} - \tfrac{1}{2}\sqrt{33} \ = \ 1\,,$$

which we may also see by inspection in (11). Using $s = 1$ and $q = -1$, we solve $-5 = 1(r - t)$ and $-6 = rt$ to get $r = -3$ and $t = 2$, and obtain the factorization

$$y^4 - 5y - 6 \ = \ (y^2 - y - 2)(y^2 + y + 3)\,.$$

Factoring the first quadratic gives $y = 2$ and $y = -1$ as solutions, and using the quadratic formula for the second quadratic gives $y = \frac{1}{2} \pm \frac{1}{2}\sqrt{-11}$ for the last two roots. Before continuing this story, I note that the first two real solutions can be found by inspection.

Back to our story. We see that the solution to linear equations is truly ancient, and that special forms of quadratic equations were known 3600 years ago, with the complete solution about 1400 years ago. Then, in the space of only a few decades about 500 years ago, cubics and quartics were also solved. By the end of the 16th century, one might have thought that with the quickening pace, a complete solution to an equation of any degree would be soon discovered.

That was not the case, and despite much work, there was no progress on this problem even for the quintic until 1799. Let me be more precise on what is meant by a solution. The solutions to equations of degree at most four expressed the roots in terms of the coefficients, where we used basic arithmetic operations, together with extracting roots. The problem of solving the quintic came to mean finding an expression for the roots of a quintic as a function of the coefficients using basic operations of arithmetic, together with extracting roots.

In 1799, Ruffini published an incomplete proof of the impossibility of this problem, and the tragically great Norwegian mathematician Niels Hendrik Abel published a complete proof of this in 1824 (when he was 22 years old). A few years later this impossibility was better understood in the 1830 treatise of (the equally tragic) Évariste Galois (who was only 18 at the time).

Here are some cubics to try your hand at solving:

(1) $y^3 - 7y - 6 = 0$.

(2) $y^3 + 2y + 4\sqrt{-1} = 0$.
(3) $y^3 - 6y - 6 = 0$.
(4) $2x^3 - 30x^2 + 162x - 350 = 0$.
(5) (Bonus!) $x^4 - 10x^2 + 4x + 8 = 0$.