

Everything you “wanted” to know about cubic equations (but were afraid to ask)

We know that if you are given any quadratic equation with real coefficients, it can be rewritten in the form

$$ax^2 + bx + c = 0$$

for real constants a, b and c where $a \neq 0$. High-school algebra tells us that the number of solutions to this equation is governed by its *discriminant* $\Delta = b^2 - 4ac$. In particular, if $\Delta < 0$, then the equation has no (real) solutions. If $\Delta = 0$, the equation has one (repeated solution), namely $x = -b/2a$. If $\Delta > 0$, then the equation has two solutions, namely

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

This is the famous quadratic formula, known to humanity in some form or another at least as far back as the 600s AD.

Here is the question: is there an analogous formula for cubic equations? First of all, let's talk about how you can simplify the situation. Given a cubic equation, by combining like terms it can be rewritten as

$$ax^3 + bx^2 + cx + d = 0$$

Divide through the entire equation by a and then perform the following substitution: let $t = x + \frac{b}{3a}$ (this is analogous to “completing the square” in a quadratic equation). Then $x = t - \frac{b}{3a}$; substituting this into the cubic equation we obtain

$$\left(t - \frac{b}{3a}\right)^3 + \frac{b}{a}\left(t - \frac{b}{3a}\right)^2 + \frac{c}{a}\left(t - \frac{b}{3a}\right) + \frac{d}{a} = 0.$$

Multiply out and combine like terms to obtain

$$t^3 + \left(\frac{3ac - b^2}{3a^2}\right)t + \left(\frac{2b^3 - 9abc + 27a^2d}{27a^3}\right) = 0.$$

In particular, there is no t^2 term in this resultant cubic. So after this substitution, we need only to consider cubic equations of the form

$$t^3 + pt + q = 0$$

where p and q are real constants (for the most general cubic $ax^3 + bx^2 + cx + d = 0$, $p = \frac{3ac - b^2}{3a^2}$ and $q = \frac{2b^3 - 9abc + 27a^2d}{27a^3}$).

There are a variety of methods to solve this cubic; here we discuss the oldest method which is due to del Ferro and Tartaglia, published by Cardano in the 1500s. The trick is to replace the single variable t with two variables u and v . Let $t = u + v$ and substitute into $t^3 + pt + q = 0$ to obtain

$$u^3 + 3u^2v + 3uv^2 + v^3 + p(u + v) + q = 0,$$

i.e.

$$u^3 + v^3 + (3uv + p)(u + v) = -q$$

From looking at this equation, we see that if u and v are two numbers satisfying $u^3 + v^3 = -q$ and $3uv + p = 0$, then $t = u + v$ is a solution of $t^3 + pt + q = 0$. But

$$\begin{cases} u^3 + v^3 = -q \\ 3uv + p = 0 \end{cases} \Leftrightarrow \begin{cases} u^3 + v^3 = -q \\ uv = -p/3 \end{cases} \Leftrightarrow \begin{cases} u^3 + v^3 = -q \\ u^3v^3 = -p^3/27 \end{cases}$$

This last system can be solved for u^3 and v^3 by substitution. We obtain

$$u^3 = \frac{-q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}} \text{ and } v^3 = \frac{-q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}$$

(which one is u and which one is v is irrelevant since u and v are interchangeable by symmetry), hence

$$t = \sqrt[3]{\frac{-q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}} + \sqrt[3]{\frac{-q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}}.$$

This solves the cubic equation $t^3 + pt + q = 0$ (sort of). Now complex numbers were not fleshed out in the 1500s, so taking the square root of a negative number was a big problem for Cardano. He solved this problem by simply assuming that $\frac{q^2}{4} + \frac{p^3}{27} > 0$. It turns out that this solution covers the cases when the cubic equation has exactly one real root, but covers no cases when the cubic has three real roots. Cardano and co. didn't know why, but we do, because we know calculus:

Consider $f(t) = t^3 + pt + q$. If $p \geq 0$, the function is everywhere increasing so it has one real root. Notice that in this case $\frac{q^2}{4} + \frac{p^3}{27} > 0$. On the other hand, suppose $p < 0$. Then the function has two critical points (one maximum and one minimum), found by setting $f'(t) = 0$, at $t = \pm\sqrt{-p/3}$. The coordinates of these points are

$$(\sqrt{-p/3}, f(\sqrt{-p/3})) \text{ and } (-\sqrt{-p/3}, f(-\sqrt{-p/3})).$$

The function has three real roots if and only if the y -coordinates of these points have opposite signs, i.e. if and only if

$$f(\sqrt{-p/3})f(-\sqrt{-p/3}) < 0.$$

If you plug in and work out what this is, this corresponds exactly to

$$\frac{q^2}{4} + \frac{p^3}{27} < 0.$$

Let $\Delta = \frac{q^2}{4} + \frac{p^3}{27}$; this is called the *discriminant* of the cubic $t^3 + pt + q = 0$. We have proven:

- $\Delta < 0$ if and only if $t^3 + pt + q = 0$ has three real roots;
- If $\Delta > 0$, the equation $t^3 + pt + q = 0$ has one real root which is

$$t = \sqrt[3]{\frac{-q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}} + \sqrt[3]{\frac{-q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}}.$$

It turns out that if $\Delta = 0$, the equation can have one or two real roots (if it has two real roots, one of them is repeated of course so you can find the repeated root by setting the derivative equal to zero and using the quadratic formula).

Remark: If you are given the general cubic $ax^3 + bx^2 + cx + d = 0$, the same discriminant is given by $\Delta = 18abcd - 4b^3d + b^2c^2 - 4ac^3 - 27a^2d^2$.

What Cardano realized, however, is that even though he didn't believe in complex numbers, if you "pretended" that negative numbers did in fact have square roots, and you "pretended" that these imaginary (Cardano called these numbers "fictitious") numbers behaved nicely algebraically, then you could manipulate the expressions for u and v so that when you added them to get t , the square roots of negative numbers disappeared. (Today, we know this is the case because if $\Delta < 0$, u and v can be taken to be complex conjugates, and when you add a complex number to its conjugate you get a real number.) In these situations you came up with one of the three real roots of the cubic.

This is why people first got interested in complex numbers. They realized that complex numbers could be used as a “tool” to find real solutions to real cubic equations. A few years after Cardano, Bombelli addressed the issues outlined in the previous paragraph and developed the laws of arithmetic for \mathbb{C} . Nonetheless, complex numbers had their detractors until the early 1800s, when Gauss and Argand used them to give a proof of the Fundamental Theorem of Algebra, which states that every polynomial with real coefficients can be factored into a product of quadratic and linear terms (equivalently, that every polynomial with complex coefficients has a complex root). Today, we know that complex numbers are not just useful in solving cubic equations, but are necessary in differential equations, quantum mechanics, relativity theory, fractal geometry, electromagnetism and fluid dynamics. They are no more imaginary than any other number.