Motivated Solution to the Cubic Equation

P. Reany

February 16, 2020

Abstract

This paper is a redo of an article that first appeared in the *Arizona Journal of Natural Philosophy* in September 1987. We show how and why to introduce a gauge into an equation to give more freedom of choice.

Generally speaking, a *motivated solution* is an attempt by an author or lecturer to present to his or her audience not only the steps to a solution, but also the following three cogonomics: 1) a brief description of the particular results being sought, 2) a brief description of the path the solution will follow, 3) an explanation of why the particular path to the goal rather than some other path. The three cogonomics can then be reduced to a minimal set of concepts that fully motivate the entire solution, and we will refer to this set of concepts as the *kernel* of the solution. For future use, we will say that a particular step is *tricky* if it is not immediately obvious, and a set of successive **obvious** steps is said to be *straightforward.*

There are a number of ways to arrive at the roots of the general cubic equation; the one we will take is McKelvey's[1] presentation of the Cardano solution. Specifically, we desire a formula for the roots to the cubic equation similar to the formula for the roots to the quadratic equation. Our starting point will be the general monic cubic equation

$$z^3 + az^2 + bz + c = 0\,, \tag{1}$$

where a, b, c are real numbers. The kernel of our method is to employ two different variable substitutions to reduce (1) to a simpler form, in particular, to a quadratic form, which we already know how to deal with. (Don't feel bad if you can't see how this is done at this point!) Our goal here is not necessarily to follow the historic approach, but rather to explain why the tricky steps to be employed are actually good rational choices.

From our experience with integration we know that well-chosen variable substitutions often transform the original integral into an easier integral. For similar reasons we will attempt variable substitutions to simplify the cubic equation. It is well know from the theory of algebraic equations that the substitution

$$z = x - a/3 \tag{2}$$

will transform (1) into the following 'reduced' cubic lacking a quadratic term.

$$x^3 \pm 3\alpha x + 2\beta = 0\,, \tag{3}$$

where

$$\pm 3\alpha = b - a^2/3 \qquad \text{and} \qquad \beta = 2a^3/27 - ab/3 + c\,. \tag{4}$$

We have used the $\pm$ sign in (3) to "factor-out" the sign, making α strictly positive for algebraic convenience later.

The next substitution, namely,

$$x = u + v \tag{5}$$

was surely a shot in the dark for whoever tried it first. This substitution into (3) gives us a single equation in two unknowns—which is not solvable without an additional constraint. By replacing one variable by two provides us an opportunities to constrain u and v (to gauge u as a function of v) to our liking to produce a system of coupled equations in both variables.

I said we needed to constrain the equation in u and v because we must now solve for u and v, and to do this we must add another equation in u and v to make two equations. This entire approach is predicated on the general principle that n simultaneous equations are needed to solve for a system of n coupled equations.

Now, using (5) in (3) we have

$$u^3 + v^3 + 3(u+v)(uv \pm \alpha) + 2\beta = 0\,. \tag{6}$$

Remember that we have already decided that it will be beneficial to mold the original equation so as to obtain an alternative equation to solve that is a quadratic, because we already know how to solve a quadratic. This quadratic could be in u or u^2 or even u^3 (the variable v enters symmetrically to u so we can deal with one without loss of generality). What we hope to gain is the ability to set a constraint on some function $f(u,v) = 0$, that is to gauge u as a function of v.

If (6) represents in our way of looking at it a constraint $f(u,v)$ on u and v, where $f(u,v) = u^3 + v^3 + 3(u+v)(uv \pm \alpha) + 2\beta = 0$, then a simple and simplifying second constraint could be

$$uv \pm \alpha = 0\,. \tag{7}$$

The advantage of this constraint is obvious, for it not only gives us a simple relations between u and v, namely,

$$u = \mp \alpha / v\,, \tag{8a}$$

$$v^3 = \mp \alpha^3 / u^3\,, \tag{8b}$$

it also simplifies (6), yielding

$$u^3 + v^3 + 2\beta = 0\,. \tag{9}$$

In equations (9) and (7) we now have two coupled equations in two unknowns. Substituting $v^3 = \mp\alpha^3/u^3$ into (9), yields

$$u^6 + 2\beta u^3 \mp \alpha = 0\,, \tag{10}$$

which is a quadratic in u^3. Therefore we obtain

$$u^3 = -\beta + \sqrt{\beta^2 \pm \alpha^3} \tag{11}$$

$$v^3 = \mp\alpha^3/u^3 = \beta - \sqrt{\beta^2 \pm \alpha^3}\,. \tag{12}$$

From here on, the procedure is straightforward: we back substitute the variables until we arrive at z. It is important to note that when we take the cube root of the previous two equations to extract u and v, we introduce the three cube roots of unity

$$1, \qquad \tfrac{1}{2}(-1 + i\sqrt{3}), \qquad \tfrac{1}{2}(-1 - i\sqrt{3})\,. \tag{13}$$

If we let

$$u_1 = [-\beta + \sqrt{\beta^2 \pm \alpha^3}]^{1/3} \tag{14}$$

$$v_1 = [-\beta - \sqrt{\beta^2 \pm \alpha^3}]^{1/3}\,. \tag{15}$$

$$u_2 = \tfrac{1}{2}(-1 + i\sqrt{3})u_1\,, \tag{16}$$

$$v_2 = \tfrac{1}{2}(-1 - i\sqrt{3})v_1\,, \tag{17}$$

$$u_3 = \tfrac{1}{2}(-1 - i\sqrt{3})u_1\,, \tag{18}$$

$$v_3 = \tfrac{1}{2}(-1 + i\sqrt{3})v_1\,, \tag{19}$$

then we obtain the three roots to (3), namely,

$$x_i = u_i + v_i \qquad (i = 1, 2, 3)\,. \tag{20}$$

And the three roots to (1) follow trivially by using (2)

$$z_i = u_i + v_i - a/3 \qquad (i = 1, 2, 3)\,. \tag{21}$$

From here on, McKelvey re-expresses these roots in terms of trigonometric and hyperbolic forms, which we will not pursue here.

J. P. McKelvey, Simple transcendental expressions for the root of cubic equations, *Am. J. Phys.* **52**, 269–270 (1984).