All you wanted to know about solving cubic equations but were afraid to ask

Peter Haggstrom
www.gotohaggstrom.com
mathsatbondibeach@gmail.com

August 30, 2014

1 Introduction

The general algebraic equation of degree n is:

$$a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n = 0 \tag{1}$$

where $a_0 \neq 0$ and the a_i can be complex.

In high school you were taught how to solve:

$$ax^2 + bx + c = 0 \quad a \neq 0 \tag{2}$$

for a, b, c real. In some cases you may have learned how to solve (2) when the coefficients were complex.

It was Abel who showed that (1) could not be solved for degree greater than or equal to 5 by the method of radicals ie in terms of the coefficients by means of $+, -, \cdot, \div$, raising to natural powers $(0, 1, 2, \dots)$ and extraction of a root of natural degree. Actually it was Ruffini who apparently proved the result before Abel but his proof was 500 pages long and it is doubtful that it was as well understood as Abel's much shorter proof.

In most high school courses the world stops at solving (2) and the reason for this (as you will soon see) is that to develop the formula for the cubic is fiddly even though it does not really involve any horrors of principle.

2 Review of quadratics

The roots of (2) are:

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \tag{3}$$

These roots are easily obtained by completing the square in (2).

First we divide by $a \neq 0$:

$x^2 + \frac{bx}{a} + \frac{c}{a} = 0$
$(x + \frac{b}{2a})^2 - \frac{b^2}{4a^2} + \frac{c}{a} = 0$
$(x + \frac{b}{2a})^2 = \frac{b^2-4ac}{4a^2}$
$x + \frac{b}{2a} = \pm\frac{\sqrt{b^2-4ac}}{2a}$
$\therefore x = \frac{-b\pm\sqrt{b^2-4ac}}{2a}$

If we let:

$$p = \frac{b}{a} \quad , \quad q = \frac{c}{a} \tag{4}$$

we get what is known as the reduced equation:

$$x^2 + px + q = 0 \tag{5}$$

Thus the roots of (5) become:

$$x_{1,2} = \frac{-p}{2} \pm \sqrt{\frac{p^2}{4} - q} \tag{6}$$

There are some basic quadratic facts that we need to recall:

1. If $\frac{p^2}{4} - q < 0$ then we get complex roots.
2. If $z \in C, z \neq 0$ and n is a natural number, there are exactly n complex numbers w such that $w^n = z$ ie there are n roots of the n^{th} degree of z.
3. x_1 and x_2 are roots of $x^2 + px + q = 0$ if and only if $x_1 + x_2 = -p$ and $x_1 \cdot x_2 = q$. This is such an important relationship it is worth proving.

If x_1 abd x_2 are roots of (5) then from (6):

$x_1 + x_2 = \frac{-p}{2} + \sqrt{\frac{p^2}{4} - q} + \frac{-p}{2} - \sqrt{\frac{p^2}{4} - q} = -p$
$x_1 \cdot x_2 = (\frac{-p}{2} + \sqrt{\frac{p^2}{4} - q})(\frac{-p}{2} - \sqrt{\frac{p^2}{4} - q}) = \frac{p^2}{4} - (\frac{p^2}{4} - q) = q$

Conversely, if $x_1 + x_2 = -p$ and $x_1 \cdot x_2 = q$ then substituting in (5) we get:

$x^2 - (x_1 + x_2)x + x_1 \cdot x_2 = (x - x_1)(x - x_2) = 0$ and so x_1 and x_2 are roots of (5).

This result is known as $Vi\grave{e}te's$ Theorem

3 The general idea of how the cubic equation is solved

Like many things in mathematics, the solution to the cubic equation involves some very simple ideas coupled with a bit of insight.

In essence one has to get from the general form of the cubic $a_0x^3 + a_1x^2 + a_2x + a_3 = 0$ to something involving a quadratic because the solution to that is known. This is a classic problem reduction approach - take something complicated, make a transformation (usually a simple one like a linear transformation) which reduces the original problem to something that has a recognisable solution. Finally, work back from the solution to the simplified problem to the original problem.

Applying this broad logic to the cubic the general approach is to:

1. "Normalise" the original equation by dividing by the leading coefficient if it is not 1.
2. Make an appropriate linear transformation.
3. Use $Vi\grave{e}te's$ Theorem appropriately. In fact this is the critical step in the whole process and is best understood by looking at what the final result is. The goal is Cardano's formula which looks like this:
$y_{1,2,3} = \sqrt[3]{-\frac{q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}} + \sqrt[3]{-\frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}}$
Note that both of the terms under the cube roots look suspiciously like they have come from a quadratic which is indeed the case.
4. Eventually take cube roots of a certain expression and then properly interpret it to work back to the original problem.

4 The detailed recipe

4.1 Normalisation

$a_0x^3 + a_1x^2 + a_2x + a_3 = 0$
$x^3 + ax^2 + bx + c = 0 \quad a = \frac{a_1}{a_0}, b = \frac{a_2}{a_0}, c = \frac{a_3}{a_0}$

4.2 Linear transformation

$x = y + d$
$(y + d)^3 + a(y + d)^2 + b(y + d) + c = 0 \quad d$ unknown at this stage
$y^3 + (3d + a)y^2 + (3d^2 + 2ad + b)y + (d^3 + ad^2 + bd + c) = 0.$ Now get rid of the squared term.
Hence $3d + a = 0 \implies d = \frac{-a}{3}$

4.3 Reduced cubic equation

We now have the reduced equation:
$y^3 + py + q = 0$ where p and q are polynomials in a,b,c:
$p = 3d^2 + 2ad + b$
$q = d^3 + ad^2 + bd + c$

4.4 Find the roots of the reduced equation

Suppose y_0 is a root of $y^3 + py + q = 0$ and that it has the form $y_0 = \alpha + \beta$
where α and β are unknowns at this stage. This leads to:
$\alpha^3 + 3\alpha^2\beta + 3\alpha\beta^2 + \beta^3 + p\alpha + p\beta + q = 0$
$\alpha^3 + 3\alpha\beta(\alpha + \beta) + \beta^3 + p(\alpha + \beta) + q = 0$
$\alpha^3 + \beta^3 + (3\alpha\beta + p)(\alpha + \beta) + q = 0$
At this stage we need a bit of inspiration. Can we use *Viète's* theorem to get something useful out of the last line?
If $\alpha\beta = \frac{-p}{3}$ then we would have $\alpha^3 + \beta^3 + q = 0$. We would then have:
$\alpha + \beta = y_0$ (4.4.1)
$\alpha^3 + \beta^3 = -q$ (4.4.2)
$\alpha^3\beta^3 = \frac{-p^3}{27}$ (4.4.3)
This suggests that *Viète's* theorem might be a goer with the roots being α^3 and β^3.

4.5 Application of $Vi\grave{e}te's$ theorem

We know that the reduced cubic equation in (4.3) has 3 roots and we can say that one of them is y_0 which we can represent as $y_0 = \alpha + \beta$ where α and β are unknown at this stage.
By $Vi\grave{e}te's$ theorem we can say that for any such y_0 the α and β indeed exist and they will be roots of this equation:
$w^2 - y_0 w - \frac{p}{3} = 0.$ Why?
Because $w^2 -$ (sum of roots)w + product of roots $= 0$.
The point of this is that the form $y_0 = \alpha + \beta$ makes sense and is consistent with the constraints
Again using $Vi\grave{e}te's$ theorem and the information from 4.4 we can see that α^3 and β^3 are roots of the equation:
$w^2 + qw - \frac{p^3}{27} = 0$ here sum of roots $= -q$ and product of roots $= \frac{-p^3}{27}$
Because this is a quadratic we can solve it - the two roots are:
$w_1 = \alpha^3 = -\frac{q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}$ (4.5.1)
$w_2 = \beta^3 = -\frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}$ (4.5.2)
Note that the square root must be one of the defined values of that operation.
This will become clearer later.
Hence the roots of $y^3 + py + q = 0$ are:
$y_{1,2,3} = \sqrt[3]{-\frac{q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}} + \sqrt[3]{-\frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}}$ CARDANO'S FORMULA
Note that for each of the cubic roots in the first part of the formula one must take the corresponding one the second subject to the constraint $\alpha\beta = -\frac{p}{3}$ (see 4.4)

4.6 Working our way back to the original problem

We now have to work our way back to the original problem systematically.
First, we note that $x_{1,2,3} = y_{1,2,3} - \frac{a}{3}$ where $a = \frac{a_1}{a_0}$. See 4.2 and recall that $x = y + d$
We also have the following relationships which arise from the reduced cubic form - see 4.3:
$3d^2 + 2ad + b = p$ (4.6.1)
$d^3 + ad^2 + bd + c = q$ (4.6.2)
(4.6.1) becomes:
$3\frac{a^2}{9} - 2a\frac{a}{3} + b = p \implies p = \frac{3b-a^2}{3}$ (4.6.3)
(4.6.2) becomes:
$(\frac{-a}{3})^3 + a(\frac{-a}{3})^2 + b(\frac{-a}{3}) + c = q$
$\frac{-a^3}{27} + \frac{a^3}{9} - \frac{ab}{3} + c = q \implies q = \frac{27c-9ab+2a^3}{27}$ (4.6.4)
This means that:
$x_{1,2,3} = y_{1,2,3} - \frac{a_1}{3a_0}$ where:
$y_{1,2,3} = \sqrt[3]{-\frac{q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}} + \sqrt[3]{-\frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}}$ and p and q are defined in (4.6.3) and (4.6.4) respectively.

4.7 Applying the formula to an example

To see how the formula works in detail it is necessary to do an example. To this end I'll concoct an example whose analytical solution is easy:

$x^3 + 3x^2 + 3x + 1 = 27 \quad (4.7.1)$

We know that this equation can be written as:

$(x+1)^3 = 27 = 3^3 e^{2k\pi i} \quad k = 0, 1, 2$

$\therefore x_{1,2,3} = -1 + 3e^{\frac{2k\pi i}{3}} \quad k = 0, 1, 2$

$x_1 = 2 \quad (4.7.2)$

$x_2 = -1 + 3(-\frac{1}{2} + i\frac{\sqrt{3}}{2}) = -\frac{5}{2} + 3i\frac{\sqrt{3}}{2} \quad (4.7.3)$

$x_3 = -1 + 3(-\frac{1}{2} - i\frac{\sqrt{3}}{2}) = -\frac{5}{2} - 3i\frac{\sqrt{3}}{2} \quad (4.7.4)$

Does our formula work? First, we rewrite (4.7.1) as:

$x^3 + 3x^2 + 3x - 26 = 0 \quad (4.7.5)$

Then, $x_{1,2,3} = y_{1,2,3} - \frac{a_1}{3a_0} = y_{1,2,3} - \frac{3}{3.1} = y_{1,2,3} - 1 \quad (4.7.5)$

Note that:

$a = \frac{3}{1} = 3 \quad (4.7.6)$

$b = \frac{3}{1} = 3 \quad (4.7.7)$

$c = -26 \quad (4.7.7)$

$p = \frac{3b-a^2}{3} = 0 \quad (4.7.8)$

$q = \frac{27c-9ab+2a^3}{27} = \frac{27\times-26-81+2\times27}{27} = \frac{27(-26-3+2)}{27} = -27 \quad (4.7.9)$

$$y_{1,2,3} = \sqrt[3]{\underbrace{-\frac{q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}}_{\alpha^3}} + \sqrt[3]{\underbrace{-\frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}}_{\beta^3}}$$

Recall from (4.5.1) and (4.5.2) that:

$\beta^3 = \frac{27}{2} - \sqrt{\frac{27^2}{4}} = 0$

$\alpha^3 = \frac{27}{2} + \sqrt{\frac{27^2}{4}} = 27$

$\therefore \alpha = 3e^{\frac{2k\pi i}{3}} \quad k = 0, 1, 2$ which means that the solutions for α are:

$\alpha_1 = 3 \quad k = 0 \quad 4.7.10$

$\alpha_2 = 3(\cos\frac{2\pi}{3} + i\sin\frac{2\pi}{3}) = 3(-\frac{1}{2} + i\frac{\sqrt{3}}{2}) = -\frac{3}{2} + i\frac{3\sqrt{3}}{2} \quad k = 1 \quad 4.7.11$

$\alpha_3 = 3(\cos\frac{4\pi}{3} + i\sin\frac{4\pi}{3}) = 3(-\frac{1}{2} - i\frac{\sqrt{3}}{2}) = -\frac{3}{2} - i\frac{3\sqrt{3}}{2} \quad k = 1 \quad 4.7.12$

So from: (4.4.1)

$$y_{1,2,3} = \alpha + \beta = \begin{cases} 3 \\ -\frac{3}{2} + i\frac{3\sqrt{3}}{2} \\ -\frac{3}{2} - i\frac{3\sqrt{3}}{2} \end{cases}$$

Therefore, using (4.7.5):

$$x_{1,2,3} = y_{1,2,3} - 1 = \begin{cases} 2 \\ -\frac{5}{2} + i\frac{3\sqrt{3}}{2} \\ -\frac{5}{2} - i\frac{3\sqrt{3}}{2} \end{cases}$$

Thus we get the same as (4.7.2)-(4.7.4) and so the formula works.

Clearly this is a fiddly process and one has to have some admiration for Cardano in working out his formula. It also explains why in high school the solutions to cubics and quartics are not usually pursued.