The Solution of a Cubic Equation

1. Introduction

The formula to find the zeros of a quadratic equation has been known for thousands years. However, it was in the middle of 16th century that the solution of a cubic equation was generated. The person who is said to be the first to create it is Scipione del Ferro (1465-1526). He was the professor of mathematics in the University of Bologna. Though he did not publish his discovery, he told it to his student, Antonio Maria Frior.

Then Niccolò Tartaglia learned it. At that time, it was popular for mathematicians to give a difficult math problem to others and compete with each other. This is why Ferro did not show his discovery to public. At first Tartaglia also did not tell his knowledge about the cubic equation to other people. But Gerolamo Cardano (1501-1576) entreated Tartaglia to tell him it. He made a promise not to publish it. However, he published it in the book '*Ars Magna*' in 1545 together with the formula for a quadratic equation. It was solved by Cardano's pupil, Ludovico Ferrari.

2. Cubic formula (Cardano's method)

Assume that the cubic equation is in the form of (1).

$$x^3 + ax^2 + bx + c = 0 \quad (1)$$

By substituting$x = z - \gamma$, we can transform the equation to (2).

$$(z-\gamma)^3 + a(z-\gamma)^2 + b(z-\gamma) + c = 0 \quad (2)$$

$$(z^3 - 3z^2\gamma + 3z\gamma^2 - \gamma^3) + a(z^2 - 2z\gamma + \gamma^2) + b(z-\gamma) + c = 0 \quad (3)$$

$$z^3 + (-3\gamma + a)z^2 + (3\gamma^2 - 2a\gamma + b)z + (-\gamma^3 + a\gamma^2 - b\gamma + c) = 0 \quad (4)$$

To eliminate the term of x^2, we have to substitute $\gamma = \frac{1}{3}a$ and the expression changes to (5).

$$z^3 + \left(\frac{1}{3}a^2 - \frac{2}{3}a^2 + b\right)z + \left(-\frac{1}{27}a^3 + \frac{1}{9}a^3 - \frac{1}{3}ab + c\right) = 0 \quad (6)$$

$$z^3 + \left(-\frac{1}{3}a^2 + b\right)z + \left(\frac{2}{27}a^3 - \frac{1}{3}ab + c\right) = 0 \quad (7)$$

Now we use $3p = (-\frac{1}{3}a^2 + b)$ and $2q = (\frac{2}{27}a^3 - \frac{1}{3}ab + c)$ to make the expression more simple.

So the expression (7) becomes

$$z^3 + 3pz + 2q = 0 \quad (8)$$

Here we put $z = u + v \quad (u + v \neq 0)$, and it becomes

$$(u+v)^3 + p(u+v) + 2q = 0 \ \ (9)$$

$$u^3 + 3u^2v + 3uv^2 + v^3 + 3pu + 3pv + 2q = 0 \ \ (10)$$

$$u^3 + 3(p+uv)(u+v) + v^3 + 2q = 0 \ \ (11)$$

If we impose another condition[1] that $p + uv = 0$, we obtain 2 expressions which are

$$u^3 + v^3 = -2q \ \ (12)$$

$$u^3v^3 = -p^3 \quad (\because uv = -p) \ \ (13)$$

From these 2 equations, we can solve u^3 and v^3 by means of the formula of a quadratic equation. We know the sum and product of u^3 and v^3. So we can create an equation whose zeros are u^3 and v^3.

$$k^2 + 2qk - p^3 = 0 \ \ (14)$$

And the zeros are

$$u^3 = -q \pm \sqrt{q^2 + p^3} \ \ (15)$$

$$v^3 = -q \mp \sqrt{q^2 + p^3} \ \ (16)$$

In (12) and (13), we can the variables u and v because the order does not affect the sum and product. So we decide that

$$u^3 = -q + \sqrt{q^2 + p^3} \ \ (17)$$

$$v^3 = -q - \sqrt{q^2 + p^3} \ \ (18)$$

To solve the equation of (17) and (18), let's consider much more simple equation $m^3 - 1 = 0$.

If we factorize $m^3 - 1$ to be $(m-1)(m^2 + m + 1)$, we know that

$$m - 1 = 0 \quad \text{or} \quad (m^2 + m + 1) = 0$$

Therefore the zeros of this equation are

$$m_1 = 1 \qquad m_2 = \frac{-1+i\sqrt{3}}{2} \qquad m_2 = \frac{-1-i\sqrt{3}}{2}$$

[1] In this case, the number of solutions including complex zeros or multiple roots in a cubic equation must be 3. Even if I add another condition which does not seem inevitable, the result contains 3 solutions. So the insert of another condition like this does not matter.

Generally speaking, if we solve $y^3 = d^3$, the zeroes are

$$y_1 = d$$

$$y_2 = (\frac{-1 + i\sqrt{3}}{2})d$$

$$y_3 = \left(\frac{-1 - i\sqrt{3}}{2}\right)d$$

In this way, we can find 3 zeros in every cubic equation if we take complex numbers into account. Therefore the solutions of u and v are

$$u_1 = \sqrt[3]{-q + \sqrt{q^2 + p^3}}$$

$$u_2 = (\frac{-1 + i\sqrt{3}}{2})\sqrt[3]{-q + \sqrt{q^2 + p^3}}$$

$$u_3 = (\frac{-1 - i\sqrt{3}}{2})\sqrt[3]{-q + \sqrt{q^2 + p^3}} \quad (19)$$

$$v_1 = \sqrt[3]{-q - \sqrt{q^2 + p^3}}$$

$$v_2 = (\frac{-1 + i\sqrt{3}}{2})\sqrt[3]{-q - \sqrt{q^2 + p^3}}$$

$$v_3 = (\frac{-1 - i\sqrt{3}}{2})\sqrt[3]{-q - \sqrt{q^2 + p^3}} \quad (20)$$

Each variable u and v has 3 zeros, but their combinations are only 3 sets. Because we imposed the condition that $p + uv = 0$. So there are 3 values corresponding to z.

$$z_1 = u_1 + v_1$$

$$z_2 = u_2 + v_3$$

$$z_3 = u_3 + v_2 \quad (21)$$

Finally, we use $x = z - \gamma$, $\gamma = \frac{1}{3}a$, $3p = (-\frac{1}{3}a^2 + b)$, $2q = (\frac{2}{27}a^3 - \frac{1}{3}ab + c)$, expression (19), (20) and (21) to find 3 zeros of a cubic equation.

$$x_1 = z_1 - \frac{1}{3}a$$

$$x_1 = (u_1 + v_1) - \frac{1}{3}a$$

$$x_1 = \sqrt[3]{-(\frac{1}{27}a^3 - \frac{1}{6}ab + \frac{1}{2}c) + \sqrt{(\frac{1}{27}a^3 - \frac{1}{6}ab + \frac{1}{2}c)^2 + (-\frac{1}{9}a^2 + \frac{1}{3}b)^3}}$$
$$+ \sqrt[3]{-(\frac{1}{27}a^3 - \frac{1}{6}ab + \frac{1}{2}c) - \sqrt{(\frac{1}{27}a^3 - \frac{1}{6}ab + \frac{1}{2}c)^2 + (-\frac{1}{9}a^2 + \frac{1}{3}b)^3}} - \frac{1}{3}a$$

$$x_2 = z_2 - \frac{1}{3}a$$

$$x_2 = (u_2 + v_3) - \frac{1}{3}a$$

$$x_2 = (\frac{-1 + i\sqrt{3}}{2})\sqrt[3]{-(\frac{1}{27}a^3 - \frac{1}{6}ab + \frac{1}{2}c) + \sqrt{(\frac{1}{27}a^3 - \frac{1}{6}ab + \frac{1}{2}c)^2 + (-\frac{1}{9}a^2 + \frac{1}{3}b)^3}}$$
$$+ (\frac{-1 - i\sqrt{3}}{2})\sqrt[3]{-(\frac{1}{27}a^3 - \frac{1}{6}ab + \frac{1}{2}c) - \sqrt{(\frac{1}{27}a^3 - \frac{1}{6}ab + \frac{1}{2}c)^2 + (-\frac{1}{9}a^2 + \frac{1}{3}b)^3}} - \frac{1}{3}a$$

$$x_3 = z_3 - \frac{1}{3}a$$

$$x_3 = (u_3 + v_2) - \frac{1}{3}a$$

$$x_3 = (\frac{-1 - i\sqrt{3}}{2})\sqrt[3]{-(\frac{1}{27}a^3 - \frac{1}{6}ab + \frac{1}{2}c) + \sqrt{(\frac{1}{27}a^3 - \frac{1}{6}ab + \frac{1}{2}c)^2 + (-\frac{1}{9}a^2 + \frac{1}{3}b)^3}}$$
$$+ (\frac{-1 + i\sqrt{3}}{2})\sqrt[3]{-(\frac{1}{27}a^3 - \frac{1}{6}ab + \frac{1}{2}c) - \sqrt{(\frac{1}{27}a^3 - \frac{1}{6}ab + \frac{1}{2}c)^2 + (-\frac{1}{9}a^2 + \frac{1}{3}b)^3}} - \frac{1}{3}a$$

This is the Cardano's method to find 3 zeros in a cubic equation.

This is a great discovery not only because we get to be able to solve a cubic equation but also because a new concept of complex ides appeared in the process. When people solved a quadratic equation and a complex number is the solution, they just concluded that there is no answer (strictly speaking, there is no real zero). As people couldn't imagine a number whose square was negative, the concept of imagine number and complex number was not generated until Cardano's method was found.

However, if we solve a cubic equation, the complex number appears during calculation even if the zeros are all real numbers. So people had to acknowledge its idea. The solution of a cubic equation made them admit the existence of complex number and this contributed to mathematics afterwards.

3. Visual interpretation of a root

Cardano's method are the algebraic solution of a cubic eqaution. On the other hand, geometric solutions were reserached from the period of ancient Greece. Then Omar Khayyám (1048-1131), a Persian mathematician, created the geometric way to find the root of $x^3 + a^2x = b$ on coordinate plane.

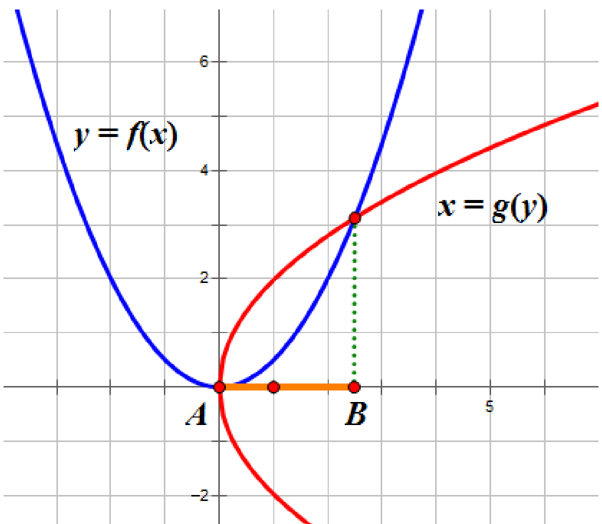


Image 1

(1) In the image 1, y=f(x) and x=g(y) are plotted.

$$f(x) = \frac{x^2}{p} \quad (1)$$

$$g(y) = \frac{y^2}{q} \quad (2)$$

And line (2) can be describes in term of x

$$y = \pm\sqrt{qx} \quad (3)$$

So the x-coordinate of intersection of these 2 lines must be the root of

$$\frac{x^2}{p} = \pm\sqrt{qx} \quad (4)$$

Square both sides and the expression becomes

$$\frac{x^4}{p^2} = qx \quad (5)$$

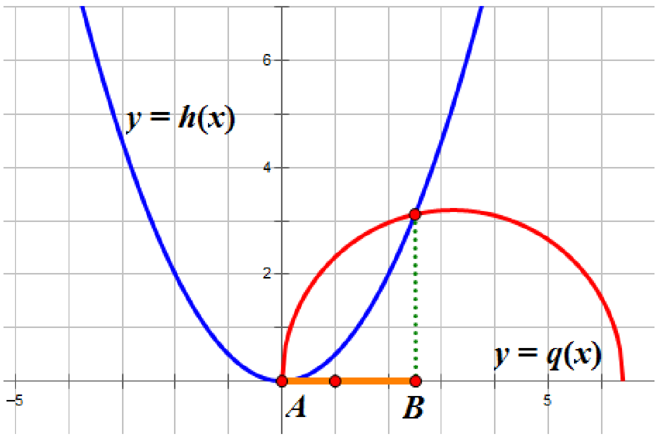
$$x^4 = p^2qx \quad (6)$$

As we are looking for the x-coordinate of intersection which is not the origin, x=0 is not what we expect as a root. Thus we can divide the equation by x.

$$x^3 = p^2q \quad (7)$$

Therefore the real root of p^2q is the x-coordinate of intersection. This corresponds to the length of A-B on the graph.

Image 2



(2) In image 2, functions h(x) and q(x) are plotted.

$$h(x) = \frac{x^2}{a} \quad (1)$$

$$q(x) = \sqrt{\frac{b^2}{4a^4} - (x - \frac{b}{2a^2})^2} \quad (2)$$

The function (2) represents a semicircle with diameter $\frac{b}{a^2}$. It passes though the origin.

We know that the common root of (1) and (2) is the x-coordinate of their intersection.

$$\frac{x^2}{a} = \sqrt{\frac{b^2}{4a^4} - (x - \frac{b}{2a^2})^2} \quad (3)$$

Square both sides and the expression becomes

$$\frac{x^4}{a^2} = \frac{b^2}{4a^4} - (x - \frac{b}{2a^2})^2 \quad (4)$$

$$\frac{x^4}{a^2} = \frac{b^2}{4a^4} - (x^2 - \frac{bx}{a^2} + \frac{b^2}{4a^4}) \quad (5)$$

$$\frac{x^4}{a^2} = -x^2 + \frac{bx}{a^2} \quad (6)$$

$$x^4 = -a^2x^2 + bx \quad (7)$$

As we already know that one solution is 0, which means 2 lines intersect at the origin, we can divide (7) by x.

$$x^3 = -a^2x + b \quad (8)$$

$$x^3 + a^2x = b \quad (9)$$

Therefore a real zero of $x^3 + a^2x = b$ is equal to the length of segment A-B.

-The exploration of the geometric solution-

Q if there are 2 or 3 real solutions in the equation $x^3 + a^2x = b$, which one will be the geometric solution?

When I was researching about this geometric solution, I noticed that if we used the coordinate plane, we can find only 1 real solution. This is why I tried to solve my question.

A. First, I have to think about the condition of equation $x^3 + a^2x - b = 0$.

1. All coefficients in an equation must be real numbers. So if there are 2 real zeros and 1 complex zero, it is impossible to have all real coefficients because –b is the product of three zeros and it cannot be a real number if there is 1 complex zero and 2 real zeros. Thus all 3 solutions must be real numbers.

2. The coefficient of x^2 must be 0.

3. The coefficient of x must be positive because a^2 cannot be negative or 0. If a^2=a=0, then $h(x) = \frac{x^2}{a}$ is undefined because a is in denominator.

4. We have to assume that b is not 0. Because the radius will be 0 and intersection will be only in the origin. Also we can factorize into $x(x^2 + a^2) = 0$ and this shows that 0 is the only real solution. Because a^2 is positive owing to 3rd condition and $(x^2 + a^2)$ cannot be 0.

Consider that zeros of equation $cx^3 + dx^2 + ex + f = 0$ $(c \neq 0)$ are α, β and γ. Due to 1st, 2nd, 3rd and 4th conditions,

$$c, d, e, f \in R \quad \wedge \alpha, \beta, \gamma \in R \quad (1)$$

$$\frac{d}{c} = 0 \quad (2)$$

$$\frac{e}{c} > 0 \quad (3)$$

$$\frac{f}{c} \neq 0 \quad (4)$$

Also we can apply the relationship between sum/product of zeros and coefficients.

$$\alpha + \beta + \gamma = -\frac{d}{c} = 0 \quad (5)$$

$$\alpha\beta\gamma = -\frac{f}{c} \quad (6)$$

$$\alpha\beta\gamma \neq 0 \quad (\because \frac{f}{c} \neq 0 \; due \; to \; 3rd \; condition\,) \quad (7)$$

So at least one of three solutions is negative and one of them is positive because of (5) and (7). None of them can be 0 because their product is not 0, and their sum must be 0. This is why we can verify it. We have already satisfied 1st, 2nd and 4th conditions. To test whether the coefficient of x can be positive or not, I have to assume that

$$\alpha < 0 \quad (8)$$

$$\beta > 0 \quad (9)$$

I can add another condition in this way because the order of solution does not matter.

I have to write the coefficient $\frac{e}{c}$ in terms of α, β, γ and it becomes

$$\frac{e}{c} = \alpha\beta + \beta\gamma + \gamma\alpha \quad (10)$$

(i) if $\gamma > 0$

$$\frac{e}{c} = \alpha\beta + \beta\gamma + \gamma\alpha \quad (11)$$

$$\frac{e}{c} = \alpha(\beta + \gamma) + \beta\gamma \quad (12)$$

$$\frac{e}{c} = (-\beta - \gamma)(\beta + \gamma) + \beta\gamma \quad (\because \alpha + \beta + \gamma = 0 \ as\ in\ (5)) \quad (13)$$

$$\frac{e}{c} = -(\beta^2 + 2\beta\gamma + \gamma^2) + \beta\gamma \quad (14)$$

$$\frac{e}{c} = -\beta^2 - \gamma^2 - \beta\gamma \quad (15)$$

As both β and γ are negative due to (9) and assumption, $-\beta^2$, $-\gamma^2$ and $-\beta\gamma$ are negative. So $\frac{e}{c}$ cannot be positive. This contradicts the 2nd condition and this shows that an equation $x^3 + a^2x - b = 0$ which satisfies all 4 conditions and has 2 positive real zeros and 1 negative zero does not exist.

(ii) if $\gamma < 0$

$$\frac{e}{c} = \alpha\beta + \beta\gamma + \gamma\alpha \quad (16)$$

$$\frac{e}{c} = \beta(\alpha + \gamma) + \alpha\gamma \quad (17)$$

$$\frac{e}{c} = (-\alpha - \gamma)(\alpha + \gamma) + \alpha\gamma \quad (\because \alpha + \beta + \gamma = 0 \ as\ in\ (5)) \quad (18)$$

$$\frac{e}{c} = -(\alpha^2 + 2\alpha\gamma + \gamma^2) + \alpha\gamma \quad (19)$$

$$\frac{e}{c} = -\alpha^2 - \gamma^2 - \alpha\gamma \quad (20)$$

Similarly, as both α and γ are negative due to (9) and assumption, $-\alpha^2$, $-\gamma^2$ and $-\alpha\gamma$ are negative. So $\frac{e}{c}$ cannot be positive. This contradicts the 2nd condition and this shows that an equation $x^3 + a^2x - b = 0$ which satisfies all 4 conditions and has 1 positive real zero and 2 negative zeros does not exist.

Therefore $x^3 + a^2x - b = 0$ cannot have 2 or 3 real zeros so that it satisfies 4 conditions. So there is not situation in which there are more than 1 real zeros in this equation.

The number of real zeros in equation $x^3 + a^2x - b = 0 \ \ (a, b \in R)$	1 real zero and 2 complex zeros	2 real zeros and 1 complex zero	3 real zeros	At least one of 3 zeros contains 0
What we can do in the geometric solution	We can find 1 real zero.	We cannot do solve an equation.	We cannot solve an equation.	We can find 0 as its zero.

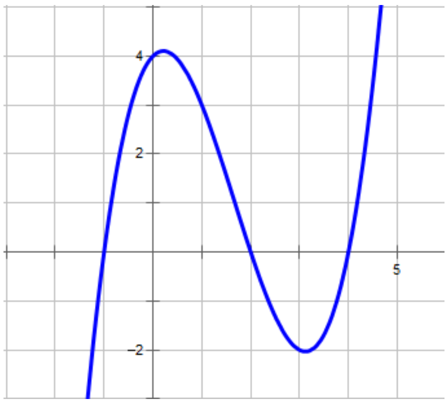
4. Discriminant of a cubic equation

The discriminant for a cubic equation $ax^3 + bx^2 + cx + d = 0 \ (a \neq 0, a, b, c, d \in R)$ is

$$\Delta = b^2c^2 + 18abcd - 4ac^3 - 4b^3d - 27a^2d^2$$

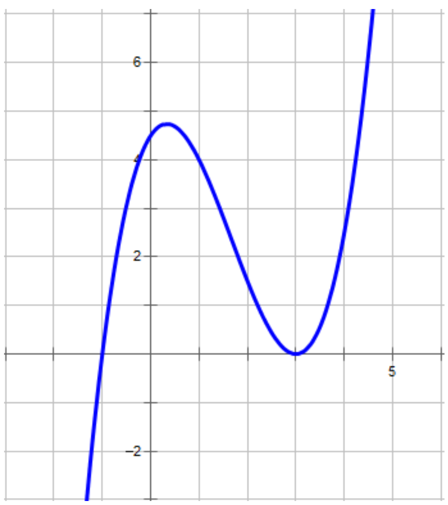
- When $\Delta > 0$, the equation has 3 different real zeros.
- When $\Delta = 0$, the equation has 1 real zero of multiplicity of 3, or 1 real zero and 1 zero of multiplicity of 2.
  In this case, if I assume that $\Delta' = -2b^3 + 9abc - 27a^2d$,
  The equation has 1 real zero of multiplicity of 3 if $\Delta' = 0$, otherwise the equation has 1 real zero and 1 zero of multiplicity of 1.
- When $\Delta < 0$, the equation has 1 real zero and 2 different conjugate complex solutions.

-Examples-

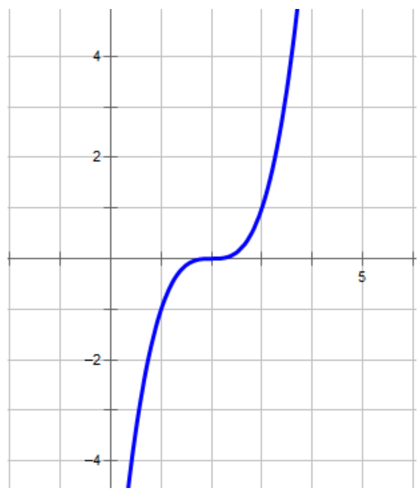


The function of this graph is $f(x) = \frac{1}{2}(x^3 - 5x^2 + 2x + 8)$. I am only interested in when $f(x) = 0$. So it does not affect the result of discriminant to stretch or shrink the function.

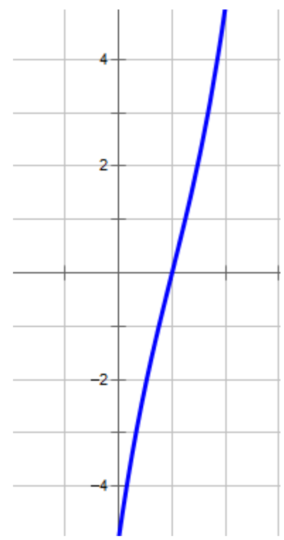
To calculate more easily, I change the original function to $2 \cdot f(x)$, which means $f'(x) = 2 \cdot f(x) = x^3 - 5x^2 + 2x + 8$. Δ of f'(x) is 900. So the discriminant of this function is positive and this indicates that the function has three different zeros. Indeed the exact zeros are x=-1, 2, 4 and they are all different zeros as the discriminant tells.



$\Delta = 0, \Delta' \neq 0$



$\Delta = 0, \Delta' = 0$



$\Delta < 0$

5. Reflection

I chose the solution of a cubic equation as the topic of math exploration because I heard that the formula to solve a cubic equation is so complicated that it is not useful in class. In middle and high school, we study how to solve a quadratic equation and its formula. For example, we can solve a math problem about 2 dimensional figures by means of quadratic formula. But the real world is made up of 3 dimensional spaces. We cannot always solve a problem of 3 dimensional figures without the general solution of a cubic equation. Compared with a quadratic equation, it took very long time for mathematicians to discover a general solution for a cubic one. Its discovery was epoch-making and this led to a new period of mathematics such as Galois Theory, which is the base of algebra. This is why I was interested in this topic.

When I learned how to solve a quadratic equation, one way to visualize its solution was to use a rectangle and transform it to a square without changing the area. So I expected that using a rectangular parallelepiped and a cube would be one of approaches to a general solution. But Cardano's method told me that a cubic equation can be solved without using geometry. The most interesting for me in Cardano's method is substitution to make the equation simpler and transform it to a quadratic form. He used what he already knew to solve a cubic equation. It was not until I explored a cubic equation that I learned how important and how useful substitutions are as a method to solve an equation. In fact, Cardano's method is nothing but the repetition of substitution. Indeed, substitution played an important role to develop algebra.

6. Conclusion

The key point in Cardano's method is substitution of the sum or difference of 2 other variables. For example, $x = z - \gamma$ and $z = u + v$. Its purpose is to eliminate term of x^2 in a cubic equation and to transform the equation to a quadratic form. So substitution plays an important role in solving a cubic equation.

In geometric solution of equation: $x^3 + a^2x = b$, its conditions are limited. We can solve it only when the zeros contain 1 real zero and 2 complex zeros. Geometric solution uses 3 types of quadratic graphs: $y = cx^2$, $x = dy^2$ and $(x - h)^2 + y^2 = h^2$. This means that quadratic graphs can generate a solution of a cubic equation in form: $x^3 + a^2x = b$.

7. References

- http://en.wikipedia.org/wiki/Cubic_equation : Wikipedia, Cubic Equation
- http://mathworld.wolfram.com/CubicFormula.html : Wolfram MathWorld, cubic formula
- http://hooktail.sub.jp/algebra/CubicEquation/ : The history of discovery of solution to a cubic equation