The Origin of The Complex Numbers

Adarsh Kishore

Contents

1 Introduction **3**

2 Why not quadratic? **4**

3 What then? **5**

4 Making things a bit easier **5**

5 The Solution **6**

6 A Simple Example **8**

7 The Cubic Discriminant **9**

8 A Compelling Example **10**

9 Bombelli's Wild Thought **11**

10 Conclusion **12**

1 Introduction

Most books on complex numbers open with the following paragraph–

Complex numbers first arose in connection with solving the quadratic equation $x^2 + 1 = 0$.

$$\begin{aligned} x^2 + 1 &= 0 \\ \implies x^2 &= -1 \end{aligned}$$

It can be seen that no real number can be the solution to this equation, because the square of a real number is always positive.

Thus, to get a solution, mathematicians created the number i*, the imaginary unit, having the property that*

$$i^2 = -1$$

And after that its maths as usual. Happy ending.

This fairy tale of quadratic equations and imaginary numbers is quite *imaginary* and non-sense, unlike the imaginary numbers themselves. The quadratic equations were certainly not what compelled mathematicians to take i seriously.

2 Why not quadratic?

Mathematics in the 16th and 17th century was still very much linked to the Greek mathematics. The Greeks were of the view that geometry is supreme, and that equations are a mere vehicle to solve geometric problems.
Consider the quadratic equation

$$x^2 = mx + c \tag{2.1}$$

The mathematicians of the time certainly knew the quadratic formula, or its equivalent form for (2.1)

$$x = \frac{m \pm \sqrt{m^2 + 4c}}{2} \tag{2.2}$$

The reason I wrote out the equation as (2.1) and not in the standard form can be explained by observing (2.2). Since Greeks held geometry supreme, for them lengths and angles were the most important. And since length cannot be negative they were not in favour of using even the *negative* numbers!

That is why in their equations the mathematicians tried to avoid the minus sign as much as possible. It can be seen in (2.2) that the only place where it makes an appearance is the (quite unavoidable) plus-minus symbol.

And then as I've said, they regarded equations as mere tools for solving geometric problems. For example, (2.1) can be thought of as the solution for the points of intersection of the parabola $y = x^2$ and the straight line $y = mx+c$. In this case, the term $m^2 + 4c$ in (2.2) is quite clear. When it is positive, roots are real and distinct (the line is secant to the parabola), when 0, it is a tangent, and when negative then the line and parabola do not intersect. The presence of impossible numbers quite rightly justifies that there is no intersection.

So then, certainly this will not make them introspect on imaginary numbers.

3 What then?

It was certainly not the quadratic equation which was the reason for the birth of complex numbers, it was the *cubic*.

Now, I'll use modern terminology. The general cubic equation is written as–

$$f(x) = ax^3 + bx^2 + cx + d \tag{3.1}$$

To see why, we'll take an interesting detour from the current topic: we'll find a solution for the cubic equation, just like we have one for quadratics.

4 Making things a bit easier

The solution of the general cubic is quite formidable, and quite unwieldy to use also. Thus we'll see if we can make things easier.

Starting with (3.1), let's make the substitution $x = X + h$, i.e., a translation of the plane. Then,

$$\begin{aligned} f(X+h) &= a(X+h)^3 + b(X+h)^2 + c(X+h) + d \\ &= aX^3 + (3ah+b)X^2 + (3ah^2+2bh)X + (ah^3+bh^2+ch+d) \end{aligned}$$

Indeed, it can be seen that by choosing $h = -\frac{b}{3a}$, we can make things easier. The term containing X^2 vanishes completely.

Thus,

$$f\left(X - \frac{b}{3a}\right) = aX^3 + \left(3a\left(\frac{b^2}{9a^2}\right) - \frac{2b^2}{3a}\right)X + \left(-\frac{b^3}{27a^2} + \frac{b^3}{9a^2} - \frac{bc}{3a} + d\right)$$

Further, by dividing out the a on both sides, we finally end up with something looking like this–

$$f(x) = x^3 + px + q \tag{4.1}$$

Such an equation, in which the cubic has the x^2 term missing, is called the **depressed cubic** (This has got nothing to do with its mental state). So what we have seen is that solving for (3.1) is equivalent to solving for (4.1). That's *something*!

5 The Solution

We are required to solve–

$$x^3 + px + q = 0 \tag{5.1}$$

Let's make the substitution $x = u + v$. We'll see the use of doing this in just a moment.

$$\begin{aligned} x &= u + v \\ \implies x^3 &= (u+v)^3 \\ \implies x^3 &= u^3 + v^3 + 3uv(u+v) \\ \implies x^3 &= u^3 + v^3 + 3uvx && \text{(Since } u + v = x\text{)} \end{aligned}$$

Now collecting the like terms together, we get

$$x^3 - 3uvx - (u^3 + v^3) = 0 \tag{5.2}$$

Now, (5.1) and (5.2) represent *identical* equations, thus their coefficients must match. Therefore,

$$3uv = -p \tag{5.3}$$

$$u^3 + v^3 = -q \tag{5.4}$$

Substituting the value of v from (5.3) into (5.4),

$$\begin{aligned} u^3 - \left(\frac{p^3}{27u^3}\right) &= -q \\ \implies 27u^6 + 27qu^3 - p^3 &= 0 \end{aligned}$$

Which is a quadratic equation in u^3. Thus we can write its solution, namely,

$$\begin{aligned} u^3 &= \frac{-27q \pm \sqrt{(27q)^2 + 4 \times 27p^3}}{54} \\ &= -\frac{q}{2} \pm \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3} \end{aligned}$$

We observe from (5.4) that u and v are symmetric, i.e., if one is taken with $+$ sign, then other will have $-$ sign. Since $x = u + v$, we have the cubic formula in all its awesomeness and glory:

$$x = \sqrt[3]{-\frac{q}{2} + \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}} + \sqrt[3]{-\frac{q}{2} - \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}} \quad ! \tag{5.5}$$

By the way, I like to call the term $\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3$ the **cubic discriminant** in analogy with the quadratic discriminant $b^2 - 4ac$ (Though this isn't standard terminology). Later on, we'll see why this is an apt name.

Thus, $D = \left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3$, and equation (5.5) becomes

$$x = \sqrt[3]{-\frac{q}{2} + \sqrt{D}} + \sqrt[3]{-\frac{q}{2} - \sqrt{D}} \tag{5.6}$$

which certainly looks more elegant!
Let's see with an example.

6 A Simple Example

Let's solve

$$x^3 - 15x - 126 = 0$$

An obvious root is $x = 6$. Try it out and you'll see.

$$\begin{aligned} D &= (63)^2 + (5)^3 \\ &= 3844 \end{aligned}$$

Thus $\sqrt{D} = 62$.
Therefore,

$$\begin{aligned} x &= \sqrt[3]{63 + 62} + \sqrt[3]{63 - 62} \\ &= 5 + 1 = 6 \end{aligned}$$

But wait. Almost immediately we notice something suspicious here. A cubic equation has three roots, but equation (5.5) apparently has just one. So where are the other two? (Hint: A number has three cube roots, and they are not all real).

If we factorize this out,

$$x^3 - 15x - 126 = (x - 6)(x^2 + 6x + 21)$$

The quadratic equation cannot be factorised, and as far as 17th century mathematics is concerned, they don't care anyway about complex numbers and the formula does its job.

But we are not satisfied, are we? I'll return to this later.

7 The Cubic Discriminant

The object of our study is the function–

$$f(x) = x^3 + px + q \tag{7.1}$$

On differentiating it once with respect to x, we obtain–

$$f'(x) = 3x^2 + p \tag{7.2}$$

(7.2) is a quadratic equation, and has two roots. Let the two roots be α and β. Then,

$$\begin{aligned} \alpha + \beta &= 0 \\ \alpha\beta &= \frac{p}{3} \end{aligned}$$

We know that for a cubic equation to have three real roots, its maximum and minimum values should have opposite signs. Thus,

$$\begin{aligned} & f(\alpha) \cdot f(\beta) < 0 \\ \implies & (\alpha^3 + p\alpha + q)(\beta^3 + p\beta + q) < 0 \\ \implies & (\alpha\beta)^3 + p\alpha\beta(\alpha^2 + \beta^2) + q(\alpha^3 + \beta^3) + pq(\alpha + \beta) + p^2(\alpha\beta) + q^2 < 0 \\ \implies & \frac{p^3}{27} + \frac{p^2}{3}\left(\frac{-2p}{3}\right) + \frac{p^3}{3} + q^2 < 0 \\ \implies & \frac{4p^3}{27} + q^2 < 0 \\ \implies & \left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3 < 0 \\ \implies & D < 0 \end{aligned}$$

Thus we can see that our cubic discriminant does indeed play a role analogous to the quadratic discriminant. In particular,

- $D < 0 \implies$ three real and distinct roots.
- $D = 0 \implies$ two real and distinct roots as two roots coincide (Actually, in case of $y = x^3$, all roots coincide at $x = 0$).
- $D > 0 \implies$ one real root and two *complex* roots.

So we are beginning to see *complex numbers*!

8 A Compelling Example

Let us try using (5.5) to solve the cubic

$$x^3 - 7x + 6 = 0$$

All its roots can be found by inspection. They are $x = 1, 2, -3$.

$$\begin{aligned} D &= \left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3 \\ &= 9 - \frac{343}{27} \\ &= -\frac{100}{27} \end{aligned}$$

Thus,

$$x = \sqrt[3]{-3 + \sqrt{-\frac{100}{27}}} + \sqrt[3]{-3 - \sqrt{-\frac{100}{27}}}$$

And the problem begins! For we have a negative term under the square root. In the case of the parabola and the straight line, the presence of negative terms under square roots in (2.2) quite rightly justified that there is no real solution. Here, it is quite apparent that the equation has not one, but three solutions!

This put the mathematicians of the time in a real fix, for they could not explain this strange solution without taking into account the fact that complex numbers exist. And one of them, Rafael Bombelli, did just this.

9 Bombelli's Wild Thought

Rafael Bombelli stumbled upon this problem while trying to use the cubic equation. He had an idea, let $\sqrt{-1}$ behave like a number in its own right! Although the symbol i may not have been used then, for our purposes we'll use it.

So he assumed that $i = \sqrt{-1}$. He further assumed that algebraically, i behaves just like normal numbers, i.e., it satisfies the commutative and the distributive law.

So if we let

$$\sqrt[3]{-3 \pm i\sqrt{\frac{100}{27}}} = \frac{1}{2} \pm ib$$

Then,

$$\begin{aligned} -3 \pm i\sqrt{\frac{100}{27}} &= \left(\frac{1}{2} \pm ib\right)^3 \\ &= \frac{1}{8} \pm i\frac{3b}{4} - \frac{3b^2}{2} \mp ib^3 \\ &= \left(\frac{1}{8} - \frac{3b^2}{2}\right) \pm i\left(\frac{3b}{4} - b^3\right) \end{aligned}$$

Comparing real part,

$$\begin{aligned} \frac{1}{8} - \frac{3b^2}{2} &= -3 \\ \implies b^2 &= \frac{25}{12} \\ \implies b &= \frac{5}{2\sqrt{3}} \end{aligned}$$

Now if we put this value of b into the imaginary part,

$$\begin{aligned} \frac{3b}{4} - b^3 &= \frac{15}{8\sqrt{3}} - \frac{125}{24\sqrt{3}} \\ &= -\frac{80`}{24\sqrt{3}} \\ &= -\sqrt{\frac{100}{27}} \quad ! \end{aligned}$$

This means that,

$$\sqrt[3]{-3 \pm i\sqrt{\frac{100}{27}}} = \frac{1}{2} \pm i\frac{5}{2\sqrt{3}}$$

and thus,

$$x = \frac{1}{2} + i\frac{5}{2\sqrt{3}} + \frac{1}{2} - i\frac{5}{2\sqrt{3}} = 1$$

Thus we get one solution. What was surprising that if we let

$$\sqrt[3]{-3 \pm i\sqrt{\frac{100}{27}}} = 1 + ic,$$
$$\sqrt[3]{-3 \pm i\sqrt{\frac{100}{27}}} = -\frac{3}{2} + id$$

then we'll get the other two also. Numbers having more than one cube roots was a new thing for the mathematicians, and it led to them taking complex numbers seriously.

10 Conclusion

Now of course, we know that if

$$z = Re^{i\theta}$$
$$\implies z^{\frac{1}{3}} = R^{\frac{1}{3}}e^{i(\frac{2\pi k}{3}+\theta)}, k = 0, 1, 2$$

but this thing caused mathematicians to first admit that the real number system was incomplete by itself, and it needed another dimension to complete it.

And *that's* the story of the origin of complex numbers. Hope you liked it!