James B. Marshall September 2019

1 Irrational numbers

Let's go back to the days of the ancient Greeks, when the concept of *number* was understood to mean either *positive whole numbers* like 1, 2, 3, etc., or *rational numbers*, which could be made from combinations (ratios) of whole numbers: $\frac{1}{2}$, $\frac{1}{3}$, $\frac{1}{4}$, $\frac{2}{3}$, $\frac{3}{4}$, $\frac{5}{2}$, $\frac{6}{5}$, etc. — and nothing else. It was perfectly clear to the Greeks that these numbers encompassed all types of quantities in existence. There was a whole number or a fraction that described every conceivable quantity in the world: a dozen eggs, two and a half apples, a third of a jug of wine, seven sixteenths of a cubit, five hundred square feet of land, and so on. What could be more obvious? The ancient Pythagoreans of the sixth century B.C. especially revered the mathematics of whole numbers, in whose orderly beauty and harmony they thought they glimpsed the mystical perfection of the gods.

But then the Pythagoreans made a revolutionary and deeply shocking discovery, which they kept secret, because they considered the knowledge too dangerous to divulge: *other kinds of numbers besides these must exist!* To see why this must be true, we first need to appreciate a couple of basic facts about whole numbers and fractions:

Fact 1: squaring a whole number always preserves even/odd-ness

Squaring an *odd* number always gives an *odd* number. Examples: $3^2 = 9$, $5^2 = 25$, $15^2 = 225$.

Squaring an even number always gives an even number. Examples: $4^2 = 16$, $6^2 = 36$, $14^2 = 196$.

This is because odd numbers do not contain any factors of 2, so the product of two odd numbers cannot contain any factors of 2 either. Conversely, since even numbers always contain at least one factor of 2, the product of two even numbers must also contain at least one factor of 2. Schematically:

 $\langle \dots$ no factor of 2 anywhere... $\rangle \times \langle \dots$ no factor of 2 anywhere... $\rangle = \langle \dots$ no factor of 2 anywhere... \rangle

 $\langle \dots$ factor of 2 somewhere... $\rangle \times \langle \dots$ factor of 2 somewhere... $\rangle = \langle \dots$ factor of 2 somewhere... \rangle

Fact 2: a fraction in lowest terms must contain at least one odd number

If a fraction $\frac{a}{b}$ is in *lowest terms*, then *a* and *b* cannot both be even. Otherwise, you could reduce it by dividing both *a* and *b* by 2. For example, $\frac{4}{6}$ reduces to $\frac{2}{3}$, and $\frac{4}{12}$ reduces to $\frac{2}{6}$, which further reduces to $\frac{1}{3}$. So if $\frac{a}{b}$ is in lowest terms (*i.e.*, it cannot be further reduced), either *a* and *b* are both odd, or one is odd and the other is even. In other words, at least one of them must be odd.

A new kind of number

Pythagoras believed that *all quantities* could be expressed as the ratio of two whole numbers in lowest terms. (A whole number itself can be written as a ratio with denominator 1, such as $\frac{5}{1}$.)

Now, consider a square of unit area, with sides equal to 1. What is the length of its diagonal?

By the Pythagorean theorem, $diagonal^2 = 1^2 + 1^2$, so $diagonal = \sqrt{2}$.

Pythagoras assumed that the square root of 2, like any quantity, must be exactly expressible as the ratio of two whole numbers a and b in lowest terms. Let's see where this assumption leads us:

$$\frac{a}{b} = \sqrt{2}$$
, with $\frac{a}{b}$ in lowest terms

so $\frac{a^2}{b^2} = 2$, by squaring both sides

so $a^2 = 2b^2$, by moving the b^2 to the right-hand side

so a^2 must be *even*, because 2 times anything gives an even number

so a itself must be even, because we know that squaring preserves even/odd-ness

that is, a = 2x for some other number x

so $a^2 = 4x^2$, by squaring both sides

so $2b^2 = 4x^2$, since we already established that $a^2 = 2b^2$

so $b^2 = 2x^2$, by dividing both sides by 2

so b^2 must be *even*, because 2 times anything gives an even number

so b itself must be *even*, because we know that squaring preserves even/odd-ness

so a and b must both be even — but this contradicts our starting assumption!

so our starting assumption *must be wrong*, because each step of the reasoning is indisputably correct, but it leads straight to a logical contradiction

that is, $\sqrt{2}$ cannot be expressed as a ratio of whole numbers!

So there *must exist* other kinds of numbers — *irrational* numbers — that express quantities such as the length of the diagonal of a simple unit square! This discovery deeply disturbed Pythagoras and his followers, because it undermined their entire worldview, which held that the mathematics of whole numbers was perfect and complete. But the logical conclusion cannot be avoided, and we are forced to expand our concept of *number* as a result.

2 Imaginary numbers

What is the solution to the equation $x^2 + 1 = 0$? Solving for x gives $x = \sqrt{-1}$, which was considered nonsensical by mathematicians of the early 16th century, who didn't believe that square roots of negative numbers could exist. They thought that such equations simply had no meaningful solutions.

The formula for solving the general quadratic equation $ax^2 + bx + c = 0$ had been known since ancient times:

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

Around 1500, the Italian mathematician Scipione del Ferro discovered an analogous but rather more complicated formula for solving the cubic equation $x^3 + px = q$, which was later rediscovered and published by Gerolamo Cardano (known as "Cardan") in 1545:

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

As an example, consider the equation $x^3 + 6x = 20$

This equation is satisfied by x = 2. Plugging p = 6 and q = 20 into Cardan's formula gives:

$$x = \sqrt[3]{\frac{20}{2} + \sqrt{\frac{20^2}{4} + \frac{6^3}{27}}} - \sqrt[3]{-\frac{20}{2} + \sqrt{\frac{20^2}{4} + \frac{6^3}{27}}}$$
$$= \sqrt[3]{10 + \sqrt{108}} - \sqrt[3]{-10 + \sqrt{108}}$$

This daunting expression works out to precisely 2, so the formula indeed gives the value we expect.

But consider another example: $x^3 - 15x = 4$

This equation is satisfied by x = 4. Plugging p = -15 and q = 4 into the formula gives:

$$x = \sqrt[3]{\frac{4}{2} + \sqrt{\frac{4^2}{4} + \frac{(-15)^3}{27}}} - \sqrt[3]{-\frac{4}{2} + \sqrt{\frac{4^2}{4} + \frac{(-15)^3}{27}}}$$
$$= \sqrt[3]{2 + \sqrt{-121}} - \sqrt[3]{-2 + \sqrt{-121}}$$

Could this crazy-looking expression containing the obviously meaningless value $\sqrt{-121}$ really be equivalent to the number 4? That made no sense at all to 16th-century mathematicians. Furthermore, the equation has two (and only two) other roots: $x = -2\pm\sqrt{3}$, and they are also indisputably real. Cardan's formula implied that $\sqrt{-121}$, whatever it was, could somehow be combined with other numbers to create ordinary real numbers. This was one of the first clues that ultimately led mathematicians to grudgingly accept the existence and legitimacy of numbers like $\sqrt{-121}$, which they nevertheless referred to, with a certain disdain, as "imaginary".

3 Euler's formula

The functions e^x , $\cos x$, and $\sin x$ can be expressed as the following infinite series:

$$e^{x} = 1 + x + \frac{x^{2}}{2!} + \frac{x^{3}}{3!} + \frac{x^{4}}{4!} + \frac{x^{5}}{5!} + \dots$$

$$\cos x = 1 - \frac{x^{2}}{2!} + \frac{x^{4}}{4!} - \frac{x^{6}}{6!} + \frac{x^{8}}{8!} - \frac{x^{10}}{10!} + \dots$$

$$\sin x = x - \frac{x^{3}}{3!} + \frac{x^{5}}{5!} - \frac{x^{7}}{7!} + \frac{x^{9}}{9!} - \frac{x^{11}}{11!} + \dots$$

Let the complex number z = a + bi. Then:

$$\begin{split} e^{z} &= e^{(a+bi)} \\ &= e^{a}e^{bi} \\ &= e^{a}\left(1 + bi + \frac{(bi)^{2}}{2!} + \frac{(bi)^{3}}{3!} + \frac{(bi)^{4}}{4!} + \frac{(bi)^{5}}{5!} + \frac{(bi)^{6}}{6!} + \frac{(bi)^{7}}{7!} + \frac{(bi)^{8}}{8!} + \frac{(bi)^{9}}{9!} + \dots\right) \\ &= e^{a}\left(1 + bi - \frac{b^{2}}{2!} - \frac{b^{3}}{3!}i + \frac{b^{4}}{4!} + \frac{b^{5}}{5!}i - \frac{b^{6}}{6!} - \frac{b^{7}}{7!}i + \frac{b^{8}}{8!} + \frac{b^{9}}{9!}i - \dots\right) \\ &= e^{a}\left(\left[1 - \frac{b^{2}}{2!} + \frac{b^{4}}{4!} - \frac{b^{6}}{6!} + \frac{b^{8}}{8!} - \dots\right] + \left[bi - \frac{b^{3}}{3!}i + \frac{b^{5}}{5!}i - \frac{b^{7}}{7!}i + \frac{b^{9}}{9!}i - \dots\right]\right) \\ &= e^{a}\left(\left[1 - \frac{b^{2}}{2!} + \frac{b^{4}}{4!} - \frac{b^{6}}{6!} + \frac{b^{8}}{8!} - \dots\right] + i\left[b - \frac{b^{3}}{3!} + \frac{b^{5}}{5!} - \frac{b^{7}}{7!} + \frac{b^{9}}{9!} - \dots\right]\right) \\ &= e^{a}\left(\cos b + i\sin b\right) \end{split}$$

This gives us a general recipe for raising e to the complex power z = a + bi:

$$e^z = e^a \left(\cos b + i \sin b\right)$$

As a special case, when $z = \theta i$, we get **Euler's formula**:

$$e^{\theta i} = e^0 \left(\cos \theta + i \sin \theta \right)$$
$$= \cos \theta + i \sin \theta$$

As another special case, when $z = \pi i$

$$e^{\pi i} = \cos \pi + i \sin \pi$$
$$= -1 + i 0$$
$$= -1$$

which can be rewritten as $e^{\pi i} + 1 = 0$, combining the five most fundamental constants in mathematics $(0, 1, \pi, e, \text{ and } i)$ using the three most important operations (addition, multiplication, and exponentiation) in a single beautiful, almost mystical, equation. It is remarkable that raising an irrational real number to an irrational imaginary power gives a negative integer as a result.

4 Complex numbers

Let z be a complex number. Think of z as a "package" of four ordinary real values: (a, b, ρ, θ) . Each of these values has a name: a is called the *real part*, b the *imaginary part*, ρ the *modulus* or *magnitude*, and θ the *phase*. These values are subject to some constraints:

- a can be anything: $-\infty < a < +\infty$
- b can be anything: $-\infty < b < +\infty$
- ρ is non-negative: $0 \le \rho < +\infty$
- θ is an angle in radians: $0 \le \theta < 2\pi$

So-called "real" numbers are just complex numbers with imaginary part 0. So-called "imaginary" numbers are just complex numbers with real part 0. Positive real numbers are just those complex numbers that have phase 0. Negative real numbers are those complex numbers with phase π . Positive imaginary numbers are complex numbers with phase $\frac{\pi}{2}$, while negative imaginary numbers have phase $\frac{3\pi}{2}$. In fact, the terms "real", "imaginary", "complex", "positive", "negative", and so on are all essentially artificial — and sometimes misleading — distinctions. As far as Nature is concerned, *all numbers are complex*, with real and imaginary parts, a magnitude, and a phase. It just took people a while (many centuries) to realize this and to come to grips with it. In some ways, the entire history of mathematics has been an ongoing struggle to understand and expand the concept of *number*, starting with simple whole numbers and progessively encompassing rational numbers, irrational numbers, zero, negative numbers, and finally imaginary and complex numbers.

Intuitively, ρ is the "size" of the number z, which in geometric terms is the distance of z from the origin in the complex plane. For numbers with phase 0 (positive reals) or phase π (negative reals), ρ is just the absolute value of the number.

Several interrelationships hold among these four values:

 $\sin \theta = \frac{b}{\rho}$ $\cos \theta = \frac{a}{\rho}$ $\tan \theta = \frac{b}{a}$ $a = \rho \cos \theta$ $b = \rho \sin \theta$ $\rho = \sqrt{a^2 + b^2}$ $\theta = \tan^{-1} \frac{b}{a}$

We can construct the number z itself from its four values (a, b, ρ, θ) in a couple of different ways. One way is to combine a, b, and i using addition and multiplication:

$$z = a + bi$$

This is sometimes called *Cartesian form*. Another way, called *exponential form*, is to combine ρ , θ , e, and i using multiplication and exponentiation:

$$z = \rho e^{\theta i}$$

This follows directly from the above interrelationships and Euler's formula:

$$z = a + bi$$

= $(\rho \cos \theta) + (\rho \sin \theta)i$
= $\rho(\cos \theta + i \sin \theta)$
= $\rho e^{\theta i}$

We can easily compute the natural (base e) logarithm of z from its ρ and θ values:

$$\log z = \log \rho e^{\theta i}$$
$$= \log \rho + \log e^{\theta i}$$
$$= \log \rho + \theta i$$

What is this formula really telling us? It says the following: Suppose you want to find the logarithm of some complex number z. For convenience, let's call that logarithm "L". To find L, you just apply the logarithm function to z's magnitude, which will give you L's real part. The phase of z simply becomes L's imaginary part. For example, take the number 3 + 4i, whose magnitude is 5 and phase is 0.927. Its logarithm L would have real part $\log(5) = 1.61$ and imaginary part 0.927. That is, L = 1.61 + 0.927i. Simple. And now that we know L's real and imaginary parts, we can easily calculate L's magnitude $\rho = \sqrt{1.61^2 + 0.927^2} = 1.86$ and phase $\theta = \tan^{-1}(\frac{0.927}{1.61}) = 0.522$. In other words, we could write L equivalently as $1.86 e^{0.522i}$.

This means that whenever we start with a positive real number, with phase 0, the logarithm L will have 0 as its imaginary part, so L will be real. That is why taking the logarithm of a positive real number always gives a real number. But negative real numbers, with phase π , and all other numbers with non-zero phase, will have logarithms with non-zero imaginary parts, so they will be complex, not real.

To square a complex number z, we just square its magnitude ρ and double its phase θ . To cube z, we cube its magnitude and triple its phase. To raise it to the 4th power, we raise ρ to the 4th power and multiply the phase θ by 4. And so on. In general: $z^n = (\rho e^{\theta i})^n = \rho^n e^{n\theta i}$.

Conversely, to find the square root of z, which is equivalent to $z^{\frac{1}{2}}$, we take the (positive) square root of its magnitude ρ and divide its phase θ by 2. To find the cube root $\sqrt[3]{z}$, we take the cube root of ρ and divide θ by 3. To find the 4th root, we compute $\sqrt[4]{\rho}$ and $\frac{\theta}{4}$. And so on. In general:

$$\sqrt[n]{z} = z^{\frac{1}{n}} = (\rho e^{\theta i})^{\frac{1}{n}} = \rho^{\frac{1}{n}} e^{\frac{\theta}{n}i} = \sqrt[n]{\rho} e^{\frac{\theta}{n}i}$$

In a moment, we'll come back to the issue of multiple roots — the fact that all numbers (except 0) always have *two* distinct square roots, *three* distinct cube roots, *four* distinct 4th roots, and so on. For now, think about what happens when we square a positive real number: since its phase is 0, doubling the phase has no effect; therefore, the result is still a positive real number. What about squaring a negative real number? Doubling its phase π gives 2π , which is equivalent to 0, modulo 2π . That is why squaring a negative real number real number.

In fact, to multiply any two numbers whatsoever, we just multiply their magnitudes ρ and *add* their phases θ (modulo 2π). Squaring, cubing, and so on are just special cases of this. That is why multiplying any two negative real numbers always gives a positive real number: the phase of their product is always 0. Likewise, multiplying two positive real numbers always gives a phase of 0 + 0, so the result is still a positive real number. Multiplying a positive imaginary number (phase $\frac{\pi}{2}$) by another positive imaginary number (phase $\frac{\pi}{2}$) gives a phase of π , so the result is a negative real number. That is why $i^2 = -1$.

How do we know that this is the right multiplication rule? That is easy to see:

$$z_1 z_2 = (\rho_1 e^{\theta_1 i})(\rho_2 e^{\theta_2 i})$$

= $\rho_1 \rho_2 e^{(\theta_1 i + \theta_2 i)}$
= $\rho_1 \rho_2 e^{(\theta_1 + \theta_2) i}$

What about multiple roots? When we raise z to the nth power, its phase θ gets multiplied by n. Conversely, when we take the nth root of z (the inverse operation), the phase gets divided by n. But because the phase is an angle in radians, it can be specified equally well by the angle $\theta + 2\pi$, or the angle $\theta + 4\pi$, or the angle $\theta + 6\pi$, or in general, the angle $\theta + 2\pi k$ for any integer $k \ge 0$, all of which are equal to the original angle θ plus some additional number of full counterclockwise revolutions around the origin.

For example, if the number z has phase $\theta = \frac{\pi}{4}$, it is equally valid to say that its phase is $\frac{9\pi}{4}$ (which is $\frac{\pi}{4} + 2\pi$), or $\frac{17\pi}{4}$ (which is $\frac{\pi}{4} + 4\pi$), or $\frac{25\pi}{4}$ (which is $\frac{\pi}{4} + 6\pi$), or $\frac{401\pi}{4}$ (which is $\frac{\pi}{4} + 100\pi$), and so on. Since these angles differ only by multiples of 2π , they all represent precisely the same phase! For that matter, so do the negative angles $\frac{-7\pi}{4}$, $\frac{-15\pi}{4}$, $\frac{-23\pi}{4}$, and so on $(\frac{\pi}{4} - 2\pi, \frac{\pi}{4} - 4\pi, \frac{\pi}{4} - 6\pi, \ldots)$. So the k in $\theta + 2\pi k$ can really be any integer at all. We usually describe θ as being in the range $0 \le \theta < 2\pi$, but this is only a convenience. What matters for the phase is just its "direction".

To find the 5th root of z, we must divide its phase θ by 5. But which angle should we use for θ ? All of the angles below are equally valid representations of z's phase, because they are all of the form $\theta + 2\pi k$ (where $k \ge 0$):

θ =		π	9π	17π	25π	33π	41π	49π	57π	65π	73π	81π	89π	97π	
	=	$\overline{4}$	4	4	4	4	4	4	4	4	4	4	4	4	• • •

If we view the above angles using "modulo 2π glasses", we see they are all really just $\frac{\pi}{4}$ in disguise:

| 0 | | π | |
|---|---|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|-----|
| θ | = | $\overline{4}$ | • • |

What happens when we divide each of the angles by 5? Here are the results:

θ		π	9π	17π	25π	33π	41π	49π	57π	65π	73π	81π	89π	97π	
$\overline{5}$	=	$\overline{20}$	$\overline{20}$	20	20	20	20	20	20	20	20	20	20	20	

Through our modulo 2π glasses, however, we see that the results are *not* all equivalent:

θ	π	9π	17π	25π	33π	π	9π	17π	25π	33π	π	9π	17π	25π	
$\frac{-}{5} =$	$\overline{20}$	$\overline{20}$	20	20	20	$\overline{20}$	$\overline{20}$	20	20	20	$\overline{20}$	$\overline{20}$	20	$\overline{20}$	• • •

In fact, there are exactly *five* distinct values, modulo 2π . Each of these corresponds to a distinct 5th-root of z. The roots all have the same magnitude $\sqrt[5]{\rho}$, but different phases, which are evenly spaced around a circle, starting at angle $\frac{\pi}{20}$ and increasing by $\frac{2\pi}{5}$ (one fifth of the circle) on each step until, five steps later, we arrive back at $\frac{\pi}{20}$.

If we take the cube root instead, dividing the phase by 3 gives three distinct new phases, modulo 2π , each one spaced a third of the way around a circle (of radius $\sqrt[3]{\rho}$) starting at $\frac{\theta}{3}$. Taking the seventh root yields *seven* distinct new phases, each one spaced one seventh of the way around starting at $\frac{\theta}{7}$. In short, numbers have multiple roots because of their phase: dividing a number's phase by n splits it into n distinct new phases, modulo 2π . In general, a number z with magnitude ρ and phase θ has exactly n distinct nth roots, each of magnitude $\sqrt[n]{\rho}$ and phase $\frac{\theta}{n} + \frac{2\pi k}{n}$, or equivalently, $\frac{1}{n}(\theta + 2\pi k)$, where $k = 0, 1, 2, \ldots, n - 1$.

This explains why positive real numbers have both a positive and a negative square root: since positive reals have phase 0, which is also equivalent to the angles 2π , 4π , 6π , 8π , etc., dividing their phase by 2, modulo 2π , yields two distinct roots, one with phase 0 (the positive root) and another with phase π (the negative root). Negative real numbers have imaginary square roots for exactly the same reasons. For example, -1 has magnitude 1 and phase π , which is also equivalent to the angles 3π , 5π , 7π , 9π , etc. Dividing the phase by 2 gives the angles $\frac{\pi}{2}$, $\frac{3\pi}{2}$, $\frac{\pi}{2}$, $\frac{3\pi}{2}$, $\frac{\pi}{2}$, \dots , modulo 2π . The numbers corresponding to these two distinct phases are *i* and -i. That is why $\sqrt{-1} = \pm i$. It all comes down to the fact that numbers have phase.

The number 1 is sometimes referred to as *unity*, and its roots are called the *nth roots of unity*. Since 1 has magnitude 1 and phase 0, its roots have magnitude 1 and phase $\frac{2\pi k}{n}$, for $k = 0, 1, 2, \ldots, n-1$. In general, the *n*th roots of unity can be expressed in $\rho e^{\theta i}$ exponential form as $e^{(2\pi k/n)i}$, for $0 \le k < n$. For example, consider the four fourth roots of unity. They are spaced evenly around the unit circle with phases in multiples of $\frac{2\pi}{4} = \frac{\pi}{2}$, namely the phases $0, \frac{\pi}{2}, \pi$, and $\frac{3\pi}{2}$. Specifically, the roots are $e^{0}, e^{(\pi/2)i}, e^{\pi i}$, and $e^{(3\pi/2)i}$, which are the numbers 1, *i*, -1, and -*i*, respectively.

Geometrically, raising each of these roots to the 4th power corresponds to starting at the root and multiplying it by itself three more times in succession, which can be visualized as counterclockwise rotations in the number plane. For root i (phase $\frac{\pi}{2}$), this gives three successive 90-degree rotations, ending at 1. For -1 (phase π), this gives three 180-degree rotations, also ending at 1. For -i (phase $\frac{3\pi}{2}$), this gives three 270-degree rotations, again ending at 1. Finally, for root 1 (phase 0), we get three 0-degree "do nothing" rotations, which simply leaves us at our starting point of 1.