Riemann Hypothesis

Introduction

The Riemann Hypothesis is the holy grail of mathematics. An unsolved problem with the potential for redefining our understanding of number theory. The Clay Institute of Mathematics has declared a 1 million USD cash prize for anyone who can solve this problem. This is some real shit.

So, lets dig a bit deeper into why this is so important. Do not worry, you do not require high level mathematics knowledge for the purposes of this article, just a bit of high school math should suffice to understand this at a very surface level. This will be a bit lengthy though, and I encourage you to skim some parts if you are already familiar with it but do so with caution.

This hypothesis has to do with the distribution of prime numbers. For the readers who are a bit rusty on primes, a prime number is one that has no factors apart from itself and 1. For example, 2 is a prime number as the only factors are 1 and 2 itself. 5 is also a prime number. 19 is also a prime number. 27 is NOT since it has factors of 1,3,9, and itself. So on and so forth.

Until now, we have not seen any pattern that describes the distribution of prime numbers along the number line. It is a mystery and prime numbers have allured mathematicians everywhere. We've been at this for almost 150 years and still no concrete pattern has been discerned. The closest we have gotten is the Riemann Hypothesis (RH). If RH was solved, then we would be able to predict where the next prime number is and how they are distributed along the number line.

Okay, let's get into a little bit of the nitty gritty of this topic.

The Riemann Hypothesis states that the non-trivial solutions for a Riemann Zeta function would all have a real part of ½.

Confused? Yeah, so was I. As promised, I will break this down piece by piece. **Remember, Riemann Hypothesis is a theory that says something about a special function called the Riemann Zeta function. Do not confuse them both.**

Funk it up with Functions

Let's have a quick discussion on what functions are for any readers who would appreciate a small detour to freshen up their knowledge. Functions are like microwaves with some settings that have been predetermined. You put something into the microwave, switch it on at whatever the predefined settings are, and then afterwards whatever you put inside is now converted to something else. A frozen pizza becomes a delicious piping hot pizza with cheese just oozing from every pore...mmmmmm...oh wait I'm getting distracted. So yeah, you put something in and something else comes out. That's a function. The inputs and corresponding outputs may change but the process undergone remains constant.

For example, lets define a function ourselves:

Function for x : $x^2 - 1$

For ease of denoting this function let's use F(x)

So, F(x) = $x^2 - 1$

Say that we input 2 into this function, then:

F(2) = $2^2 - 1$ = 4 -1 = 3

We put 2 into F(x) and 3 comes out. Amazing. Exactly like a microwave. Totally not a confusing and ridiculous analogy.

With me till now? Hopefully yes

A solution to this function F(x) would mean all possible inputs, x, such that F(x) = 0

2 CANNOT be a solution since F(2) = 3

What about F(1)?

F(1) = $1^2 - 1$ = 0

Yes! 1 is a solution of F(x).

How about -1?

F(-1) = $(-1)^2 - 1$ = 1 -1 = 0

Yes, this is also a solution of F(x).

Great! We have 2 solutions of F(x). So, this should give you an understanding of functions and solutions.

Mathematicians are a funky crowd so they decided that calling the solution a "solution" was too simple and decided to call them "zeroes" of the function instead. If you put the "zeroes" of the function into the function it becomes zero. Such genius.

So, F(x) has 2 zeroes (a.k.a. solutions): 1 and -1.

Remember this point. The zeroes of a function are those values that when inputted into a function gives out 0. Let's move on.

Quick Note:

Do you know the different sets of numbers? Natural, Whole, Integers, etc..

Natural: Integers from 1 to infinity: 1,2,3,...

Whole: Integers from 0 to infinity: 0,1,2,...

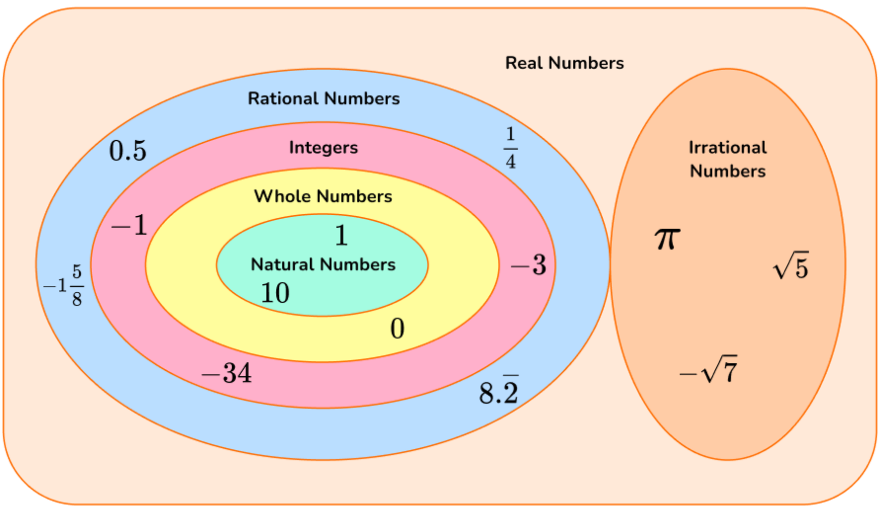
Integers: Numbers with NO decimal or fractional parts: -3,-2,-1,0,1, 2,...

Rational: All numbers that have a finite value, they can have decimal and fractional part: -56.8, ¾, 1093.22, etc..

Irrational: All numbers that do not have a finite value. Basically, when expressed in decimal form the numbers keep going and never stop. There is no definitive value they have

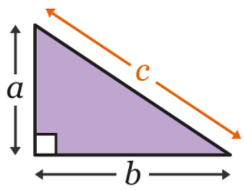
Example: π, $\sqrt{2}$, etc..

Real numbers: Rational and Irrational every possible number.



Riemann Zeta Func...what?

The Riemann Zeta function is just a kind of function. It takes in an input and gives an output after doing whacky shenanigans to the input. This is how it looks like:



$$a^2 + b^2 = c^2$$

Oh wait, this is that Pythagoras dude's theorem. Let me get the correct Riemann Zeta function.

Ah, here it is:

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$$

We represent the Riemann Zeta function with the Greek symbol ζ . The symbol ∑ means to sum up all the terms by replacing *n* with 1, then 2, then 3 and so on till infinity. The expression next to the ∑ is $1/_{n^s}$ hence we sum up $1/_{1^s}$, $1/_{2^s}$, $1/_{3^s}$ and so on and so forth. Below are a few examples:

$$\zeta(2) = \frac{1}{1^2} + \frac{1}{2^2} + \frac{1}{3^2} + \cdots till\ infinity$$

$$\zeta(3) = \frac{1}{1^3} + \frac{1}{2^3} + \frac{1}{3^3} + \cdots till\ infinity$$

If we actually try to calculate the value, using limits and some advanced maths, we get:

$$\zeta(2) = \frac{\pi^2}{6}$$

How we got this is not very important, but the point is that even though it is an infinite series, it's possible to get a value. This is possible because each term in the series gets smaller and smaller and hence, we are able to converge. Such types of infinite series where it converges to a value is called a convergent series.

What about the Riemann Zeta for -1?

$$\zeta(-1) = \frac{1}{1^{-1}} + \frac{1}{2^{-1}} + \frac{1}{3^{-1}} + \cdots till\ infinity$$

$$\zeta(-1) = 1 + 2 + 3 + \cdots till\ infinity$$

This clearly does not converge to any value, it diverges and so we say it tends to infinity. Such types of infinite series that diverge are called divergent series.

$$\zeta(-1) = \infty$$

Another interesting point here is that if we try to calculate the Riemann Zeta function of 1, that would also diverge!

$$\zeta(1) = \frac{1}{1^1} + \frac{1}{2^1} + \frac{1}{3^1} + \cdots till\ infinity$$

$$\zeta(1) = \infty$$

That seems a bit counter-intuitive since each term in the series is getting smaller but it isn't getting smaller fast enough so it still tends to infinity. It's quite weird but don't worry about understanding this fully. **What's important here is that the zeta function for values greater than 1 will always converge to a value.**

The guy who came up with this function and the hypothesis, is Bernhard Riemann. A German mathematician. The Riemann Hypothesis and the Riemann Zeta function both are his namesake. He decided to do something cheeky with the Riemann Zeta function. He decided to use complex numbers as the input...

Complexity for Simplicity's sake

I highly recommend reading through this section regardless of your familiarity. If you aren't familiar with Complex Numbers, how I envy your free and relaxed existence. It doesn't just stop at Real Numbers. There's an even bigger group.

We've been taught from a very elementary age that you CANNOT have roots of negative numbers. Guess what mathematicians did...

They created a special number, i, which is equal to $\sqrt{-1}$

Then they created a whole new set of numbers called Imaginary numbers to encompass numbers like $2i$, $3i$, -1000i, 403.7i

Then they were like hmm let's combine real numbers with imaginary numbers and gave birth to Complex Numbers: 1 + 4i, 78 + 90i, -23 + 46.8i

A complex number is a number that can be expressed as:

$$z = a + bi$$

Here, z is the complex number. a is called the Real Part and b is called the Imaginary part. This is important. Take note.

Now the big question: WHY? Why do this? Why torture our poor naïve souls with such incredulities.

The answer is that these Complex Numbers are actually amazingly useful. They are widely used in mathematics and even in the practical sciences to reduce complicated problems to simpler ones. You can now find out roots for ANY polynomial equation, and this could be used to analyze so many different things. If I asked you what the zeroes of the function g(x) = $x^2 + 3$ are, do you have an answer?

$$x^2 + 3 = 0$$

$$x^2 = -3$$

$$x = \sqrt{-3}$$

$$x = \sqrt{-1 * 3}$$

$$x = \sqrt{-1} * \sqrt{3}$$

$$x = i\sqrt{3}$$

Tada! It's not obvious why this is useful, but this is a gamechanger. From quantum mechanics to electrical engineering, complex numbers are used widely. Think of it as a helpful intermediary to reach a useful conclusion rather than the useful conclusion itself. A catalyst to help reduce complicated equations to simpler ones. You don't have to understand much about complex numbers, just that they exist and how to represent them.

Given a complex number z = 3 + 4i:

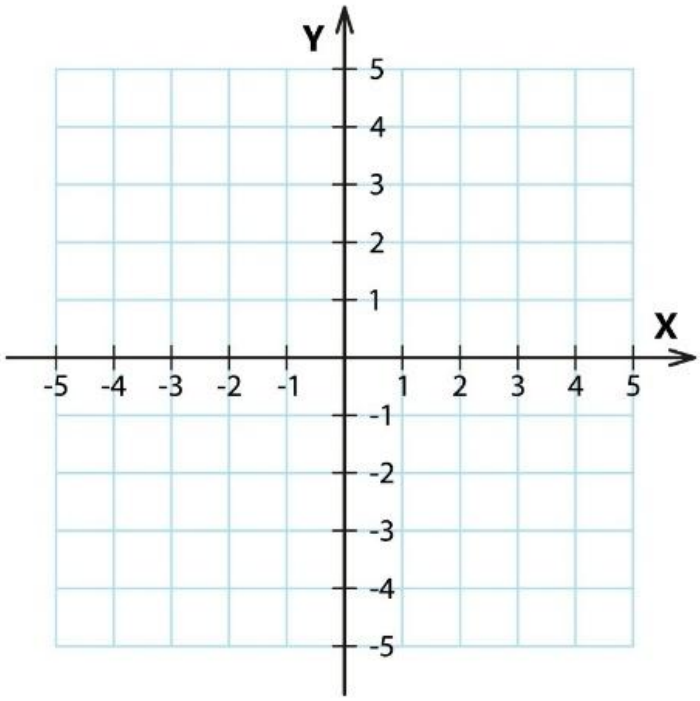
Real part of z = Re(z) = 3

Imaginary part of z = Im(z) = 4

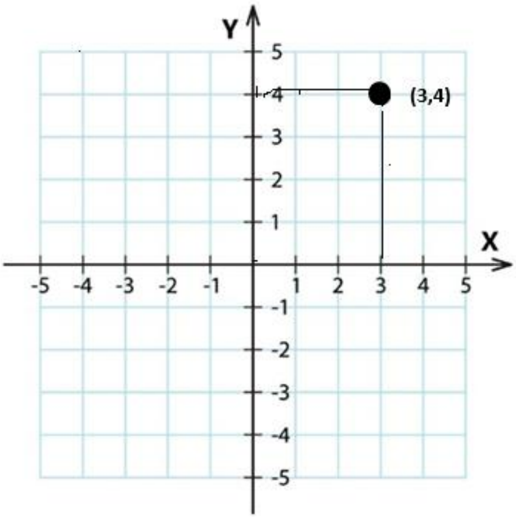
i is the square root of -1 so $i^2 = -1$

Coordinate ~~Exes~~ Axes

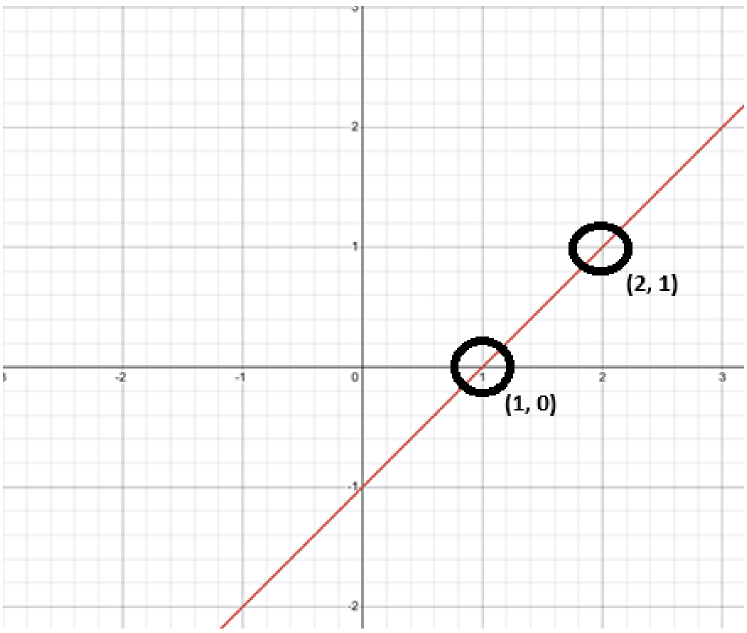
Do you remember the co-ordinate plane?



There is an x-axis and a y-axis. Each point on this plane can be represented as a point (x,y) where x is the length along x-axis and y is the length along the y-axis.



If I wanted to plot the function F(x) : x - 1 on the graph it would look like this:



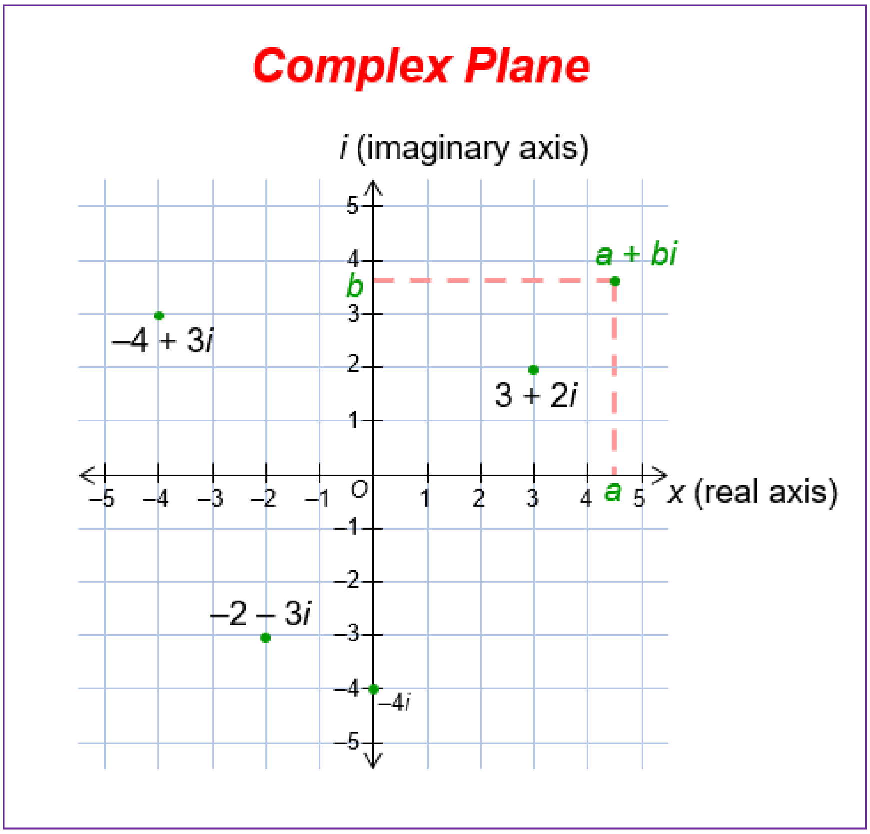
Each x-coord is the input to the function and the y-coord is the corresponding output. So, all points on the red line are (x, F(x))

If we take x = 1. Then F(1) = 1 -1 = 0. Hence when the x-coord is 1, the y coord is 0: (1, 0). This point lies on the red line. It has been circled. (2,1) is another point.

Now, say that the x-axis represents all Real numbers: -2.9, -1.99999, 1, π, 3, ...

Let's also say that the y-axis represents all Imaginary numbers: -8.6i, -1i, 1.8i, $\sqrt{2}i$, 3i...

Et voila! We get the complex co-ordinate plane and can express any complex number on this plane:



For a complex number z = a + ib

The Real part of the complex number is the x-coord. Re(z) = a

The Imaginary part of the complex number is the y-coord. Im(z) = b

Okay, so now that we have a complex plane. How do we denote a function that is in the complex plane? Just think about it for a second. Normally, we have a simple function F(x) where the input is just the x-coord and whatever the output is we will just mark that as the y-coord. For a complex plane, we need both the x and y coord to denote a specific input, so where do we mark the output? Even if your input is just an integer like 2. In the complex plane that would be denoted as 2 + *0i*. So (2,0). That uses both x and y coord to denote a single input. In the normal coordinate plane, any input is just on the x-axis directly.

A bit of a mindbender, yea?

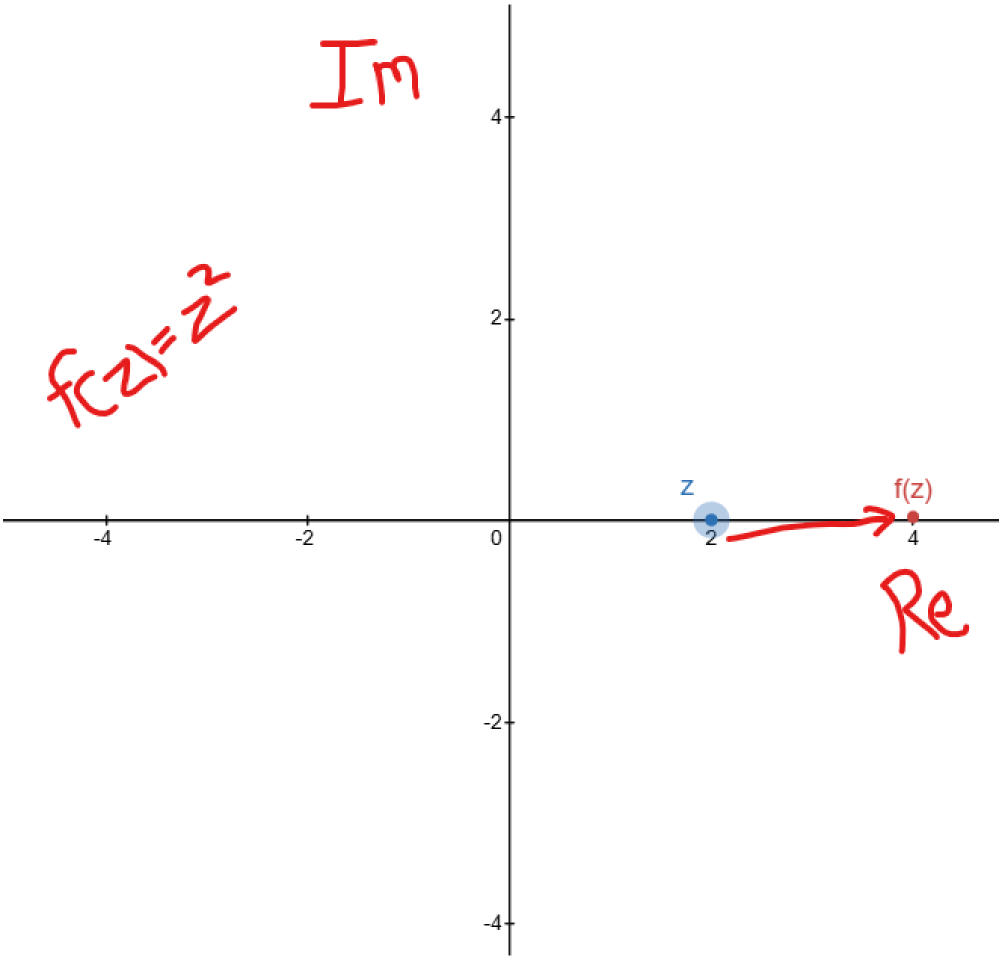
We can't really use the rules of normal coordinate systems to plot functions in the complex plane. But there's till something else we can do.

Let's take a function in the complex plane: $f(z) = z^2$ where z is a complex number.

f(2) = 2^2 = 4

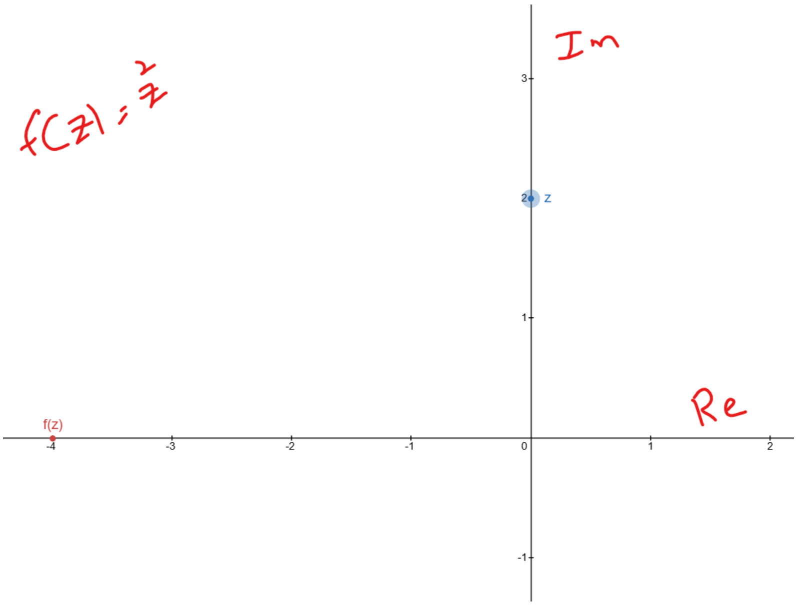
f(2i) = $(2i)^2 = 2^2 * i^2 = 4 * (-1) = -4$

On the complex plane a representation for f(2) would be:



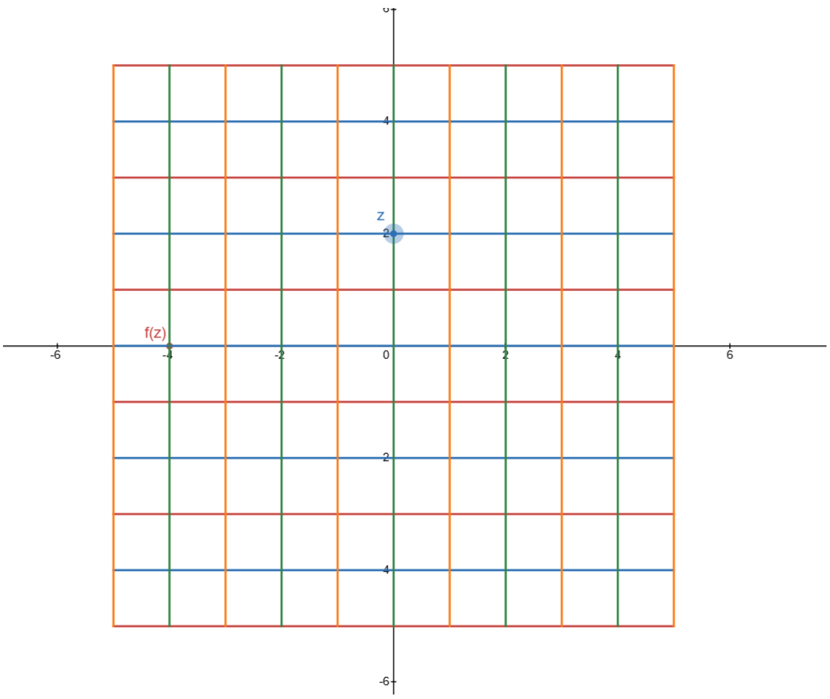
Here the blue dot represents z and the red dot represents f(z)

Similarly, f(2i) would be:

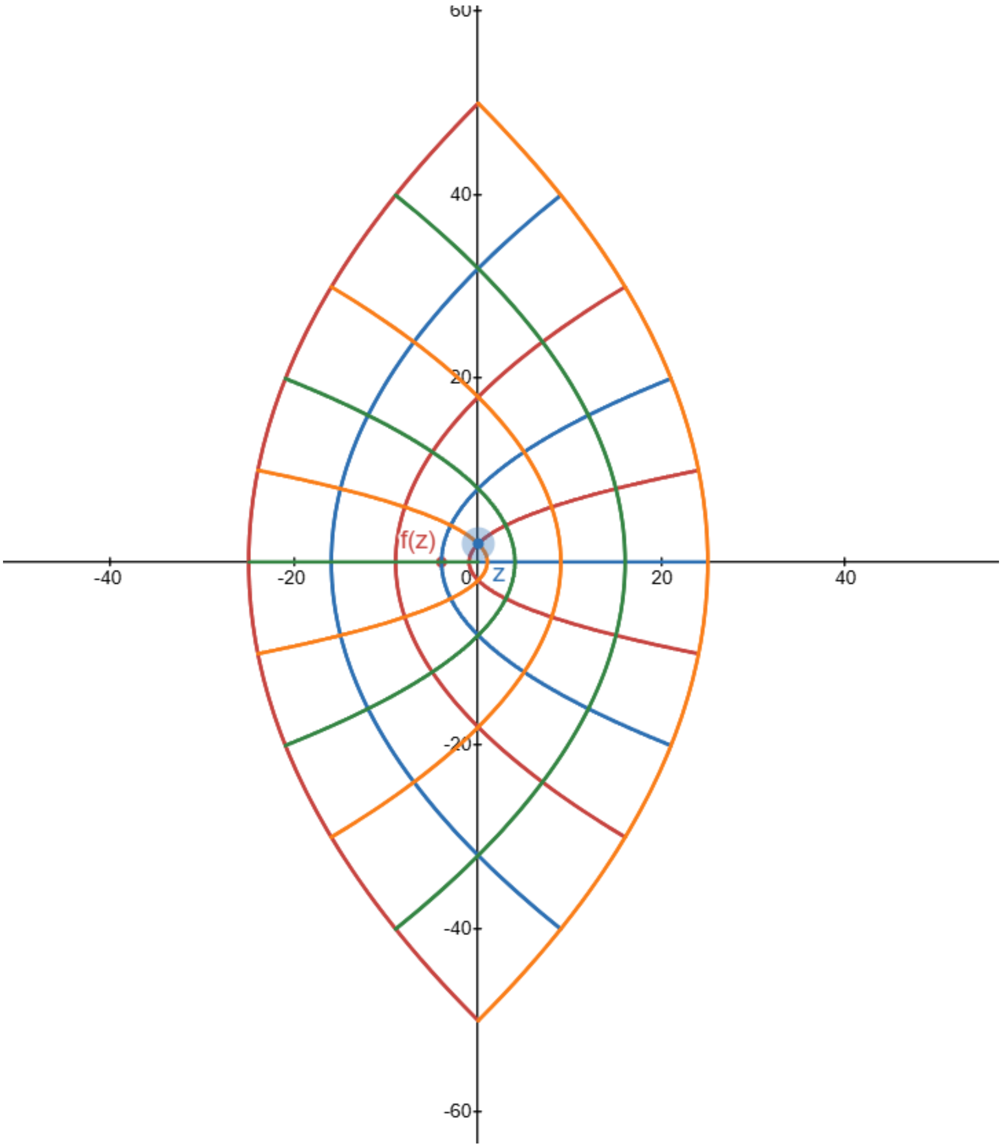


Let's introduce some gridlines. The gridlines cover the space from -5 to 5 on the Real and Imaginary axes. This is what we call our input space. We will now perform a transform of the graph by applying the function f(z) to each point in the input space and SHIFTING it to its output. Essentially, we are pulling each point to its output. You can imagine a lot of stretching and rotating that will happen. We're giving this graph a proper yoga workout.

Before transform:



After transform:



Notice how the max limit on the Real axis is 25 and on the Imaginary axis is 50. This is because the maximum Real value is from an input like (5,0) in the input space.

f(5,0) = 5^2 = 25

and the maximum imaginary value is from an input like (5,5) in the input space.

f(5,5) = $(5 + 5i)^2$ = $5^2 + (5i)^2$ = 25 + 25(-1) + 2(5)(5i) = 50i

Remember $(a + b)^2 = a^2 + b^2 + 2ab$

This is how we represent functions in the complex plane, all the possible outputs for a predefined input space are displayed. You can't really figure out which output corresponds to which input just by looking at the graph, but you will know what the possible outputs of a function are.

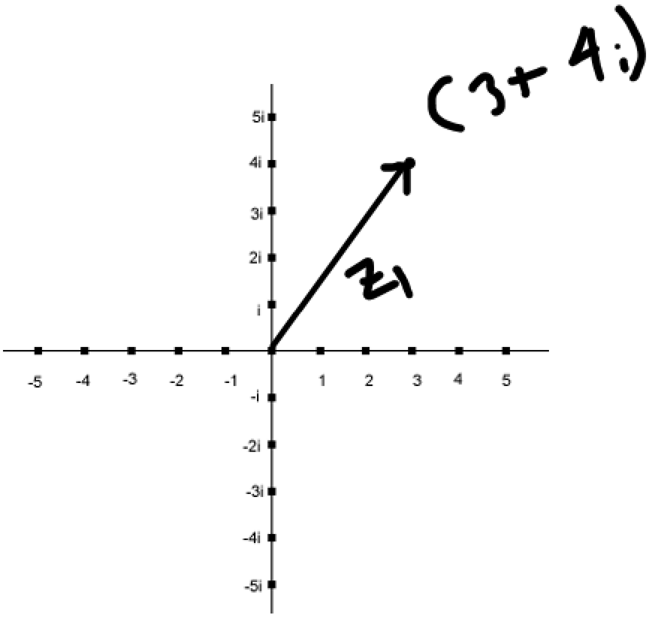
This is called complex transformation.

Complex Multiplication

Okay so we are successfully able to plot complex numbers on the graph and have even seen how to transform a function in the complex plane. Let's look at multiplication.

Take a complex number $z_1 = 3 + 4i$

Here's how it would look like on the complex plane:



What if we multiply this by $z_2 = 0 + 1i$. Just a fancy way of saying to multiply by i.

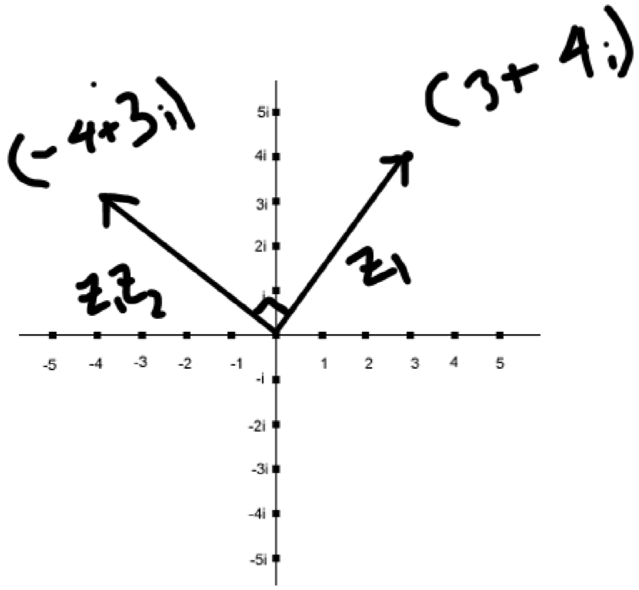
$$z_1 * z_2 = (3 + 4i) * i$$

$$z_1 * z_2 = 3i + 4i^2$$

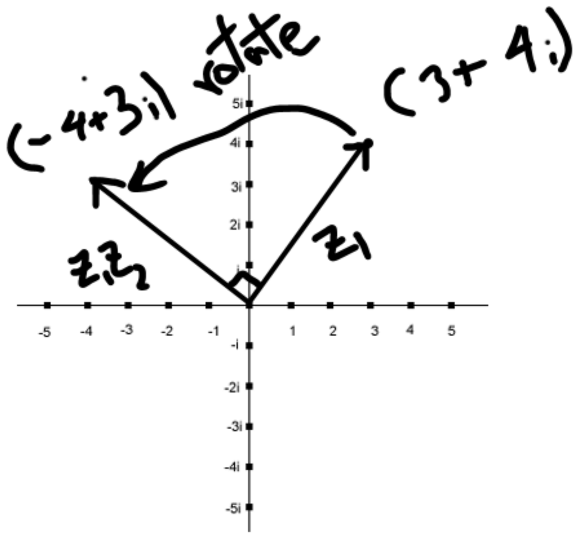
$$z_1 * z_2 = 3i + 4\,(-1)$$

$$z_1 * z_2 = -4 + 3i$$

Let's chart this on the complex plane as well



Hmm…something seems weird. We multiplied and the new point seems to be the same length from the origin, and also at a perfect 90 degrees to the original point.



Essentially, multiplying with *i* just rotated the point.

This is quite an interesting result.

So, what does this mean for the Riemann Zeta function?

Let's take a complex number z = 2 + ib. Never mind what b is.

Plugging this into the Riemann Zeta:

$$\zeta(2+\mathrm{ib}) = \frac{1}{1^{2+ib}} + \frac{1}{2^{2+ib}} + \frac{1}{3^{2+ib}} + \cdots till\ infinity$$

$$\zeta(2+\mathrm{ib}) = \frac{1}{1^2 * 1^{ib}} + \frac{1}{2^2 * 2^{ib}} + \frac{1}{3^2 * 3^{ib}} + \cdots till\ infinity$$

$$\zeta(2+\mathrm{ib}) = (\frac{1}{1^2} * \frac{1}{1^{ib}}) + (\frac{1}{2^2} * \frac{1}{2^{ib}}) + (\frac{1}{3^2} * \frac{1}{3^{ib}}) + \cdots till\ infinity$$

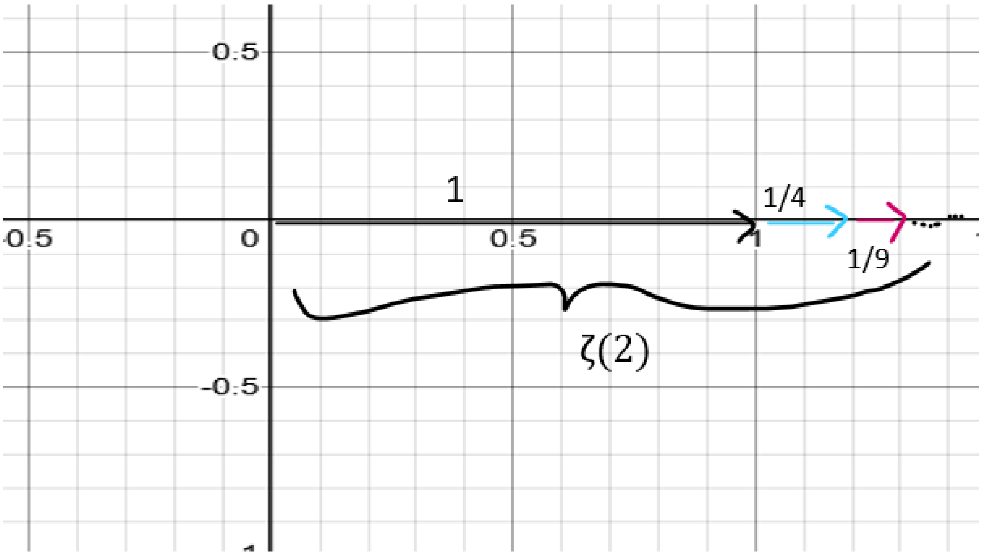
Anything raised to the power of an imaginary number, is just another complex number. And this complex number has the unique property that it does not change the magnitude at all, just rotates the point it is multiplied to! The derivation for this is out of the scope right now but it's based on Euler's formula that you should check out if you have the time. It is based on the observation we made earlier regarding the rotation when we multiply with *i*. **The important point here is that raising to the power of *i* gives us another complex number that just rotates the point it is multiplied to, nothing else.**

So now, the zeta function can be reduced to the format below:

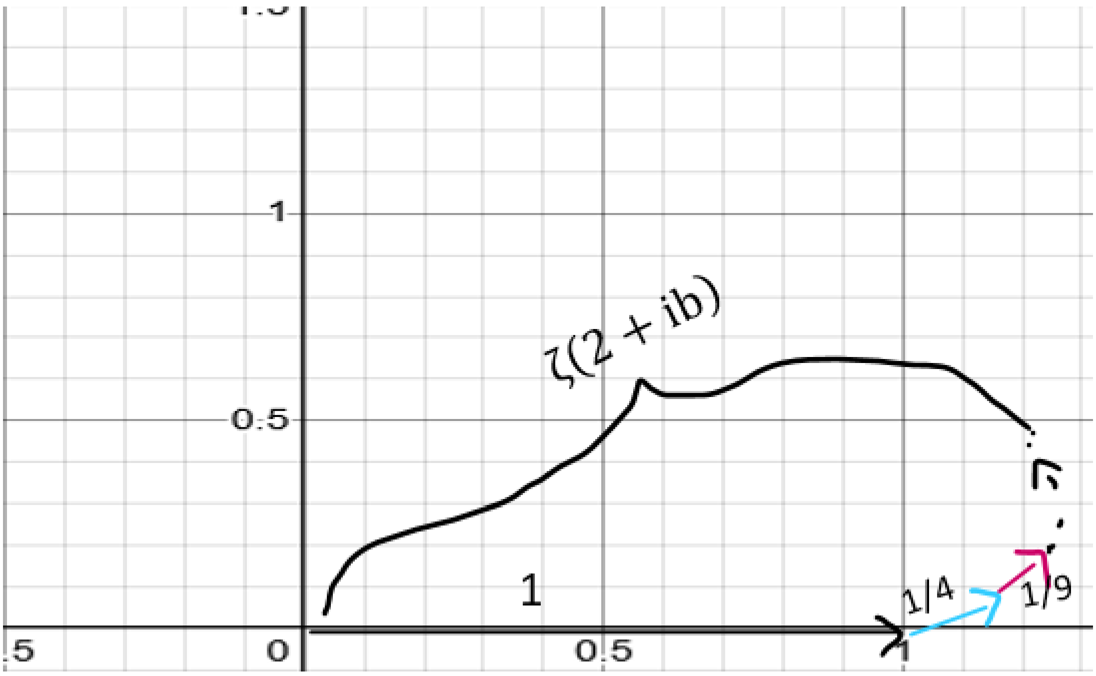
$$\zeta(2+\mathrm{ib}) = (\frac{1}{1} * z1) + (\frac{1}{4} * z2) + (\frac{1}{9} * z3) + \cdots till\ infinity$$

Normally if we were to just take the input as just 2, and try to sketch out each term it would look like this:

$$\zeta(2) = \frac{1}{1} + \frac{1}{4} + \frac{1}{9} + \cdots \textit{till infinity}$$



If we were to sketch out ζ(2 + ib) it would look like this:



The series will still converge, just at a different value since we have to also consider rotations due to imaginary values. The Imaginary part DOES NOT AFFECT the magnitude in the Riemann Zeta function. Only the rotation. The Real part affects the magnitude. This means that whether we use real numbers or complex numbers; the convergence or divergence remains the same. It is only based on the Real part. So, **even if we use complex inputs, as long as the Real part is >1 the Riemann Zeta function will converge and for all inputs with Real part <=1 it will diverge.**

We're Almost There

One final detour before we understand the Riemann Hypothesis.

The Riemann Hypothesis says that, ***the non-trivial solutions for a Riemann Zeta function would all have a real part of ½.***

Umm wait, what does non-trivial mean? Well, simply put, non-trivial is basically mathematicians saying not the obvious solutions.

For example, let's take a new function

F(x) : $x^2 - 2x$

A very obvious solution to this is if x was 0.

F(0) = 0 - 0 = 0

A not so obvious one would be 2.

F(2) = $2^2 - 2 * 2$ = 4 - 4 = 0

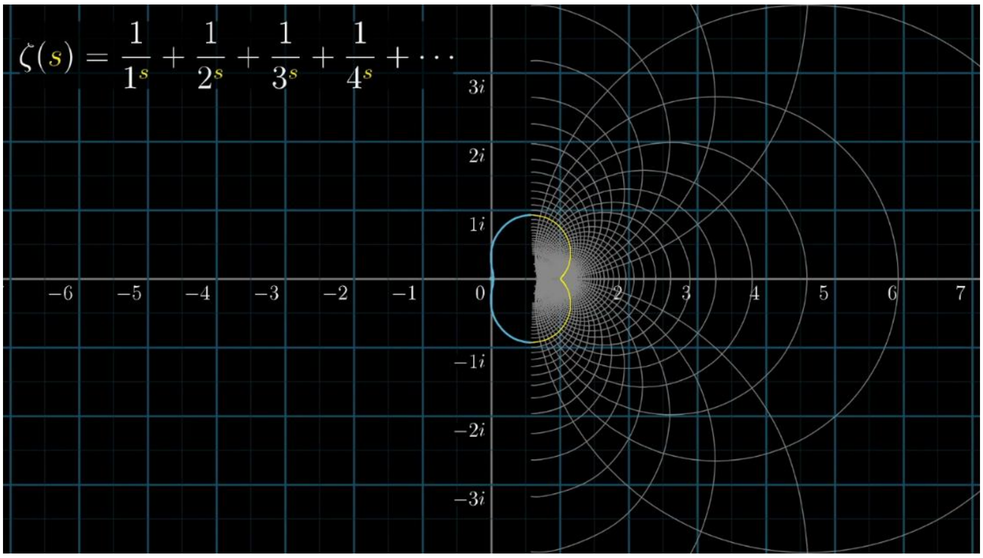
So, there are 2 zeroes(solutions) for F(x) now but only 1 is not obvious. Hence, there is one non-trivial zero for F(x).

The obvious solutions for the Riemann function are all the negative even integers like -2, -4, -6. How this came about is out of the scope of this article. Right now, it's just important you get what non-trivial means. Basically, any zero for the Riemann Zeta function that isn't a negative even integer will be considered a non-trivial solution.

Putting it all together

We have now reached a stage where we can understand the Riemann Hypothesis.

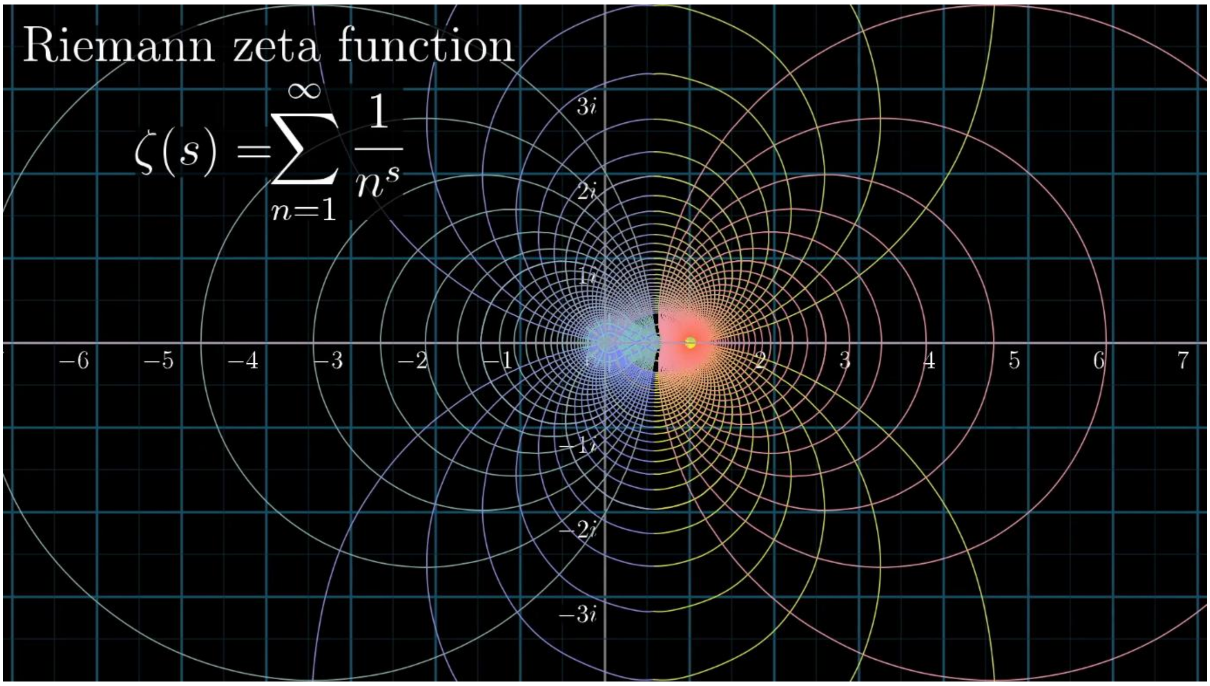
Let's graph the Rieman Zeta in the complex plane. Remember this isn't a simple graph where the input is x-coord and the output is y-coord. Each point in the graph has been shifted to its Reimann Zeta output. This means that all possible outputs of the Reimann Zeta function have been plotted:



It looks quite beautiful, right?

You will notice that a lot of the graph is blank, that's because the Riemann Zeta function is only valid for inputs whose real part is >1. Remember the convergent and divergent series discussion? If we were to input any value whose real part<= 1 then the function would become infinity. So, we have transformed the graph only for inputs with Real part > 1.

You will notice immediately that we kind of have one half of the graph. It's just begging to be completed. Let's say we want to just complete the graph for the sake of symmetry, there is actually a process called Analytical Continuation we can use. It's just completing the other half of the graph. Like making a full circle from a semi-circle. Now the function looks like this:

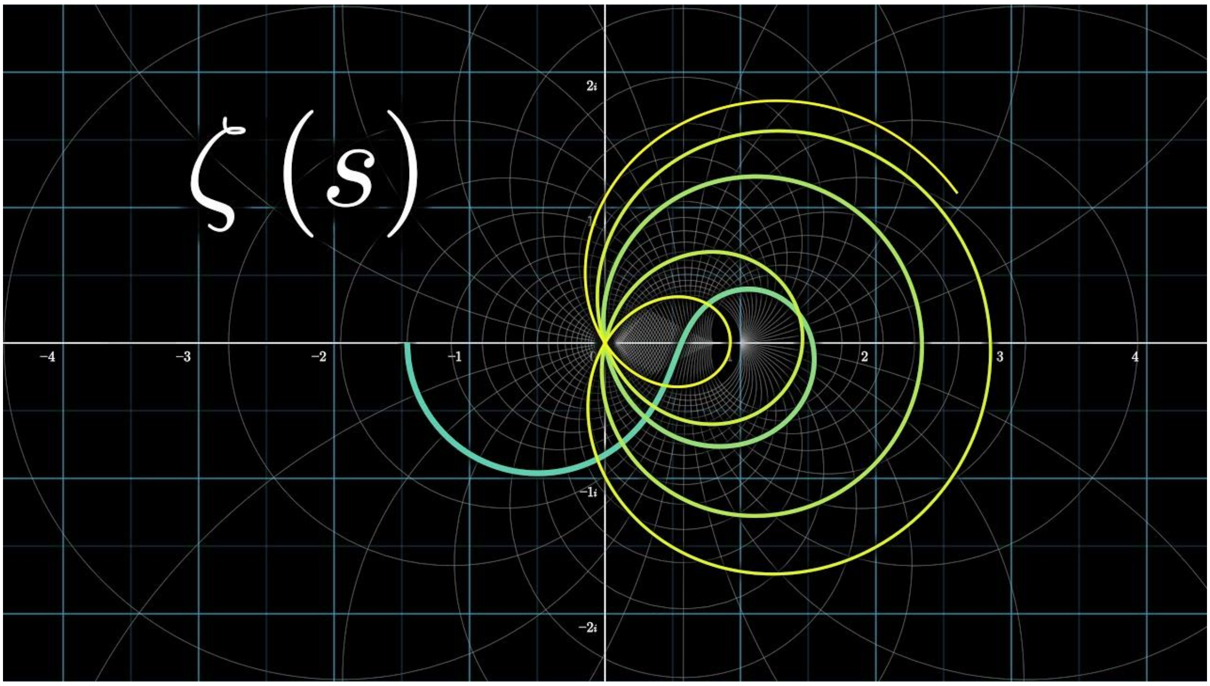


Amazing, isn't it? Truly wonderful.

This is the Riemann Zeta function in all its glory. With the applied Analytical Continuation.

Let's highlight the outputs whose inputs have real part equal to ½: So, anything in the format of $\zeta\left(\frac{1}{2} + bi\right)$ such as $\zeta(½ + 3i)$, $\zeta(½ - 7i)$, etc...

You would get the following highlighted part of the graph:



This beautiful spiral. This is what all the fuss is about. Do you see how it keeps looping through 0? The Reimann Zeta function keeps becoming 0, taking a loop and crossing 0 again. And this happens only when the real part is ½ (except at the trivial solutions at negative even integers, but we can ignore those).

So, all the non-trivial zeroes seem to only happen when the real part is ½.

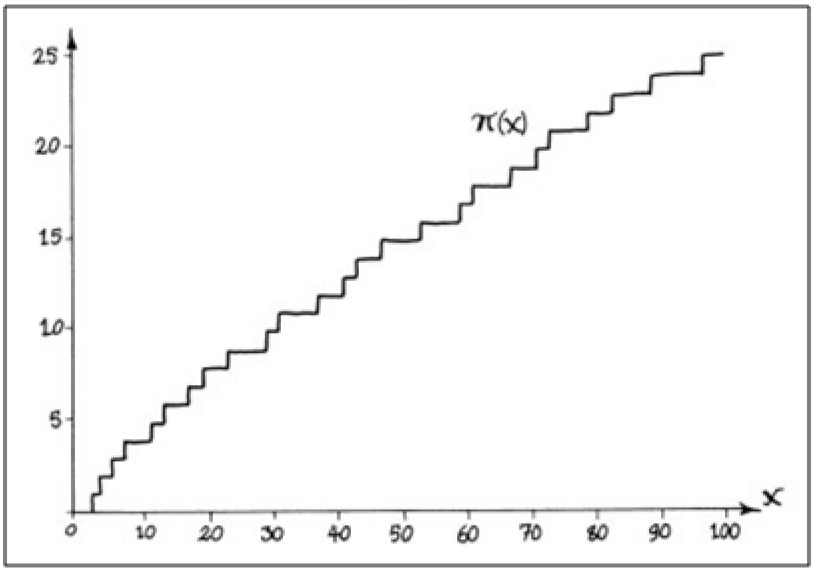
Some of the zeroes of the Riemann Zeta function are: ½ + 14.13*i* , ½ + 21.02*i* , ½ + 25.01*i*

And this my friends, is the Riemann Hypothesis. That all non-trivial zeroes have a real part of ½. There is absolutely no way we can definitely prove it as of yet, we can only keep trying to find more and more zeroes of the Riemann Zeta function and if even one of them does not have a real part of ½ then it is disproved. We have had supercomputers calculate the first 10 TRILLION solutions to this function and all of them have had the real part 1/2.

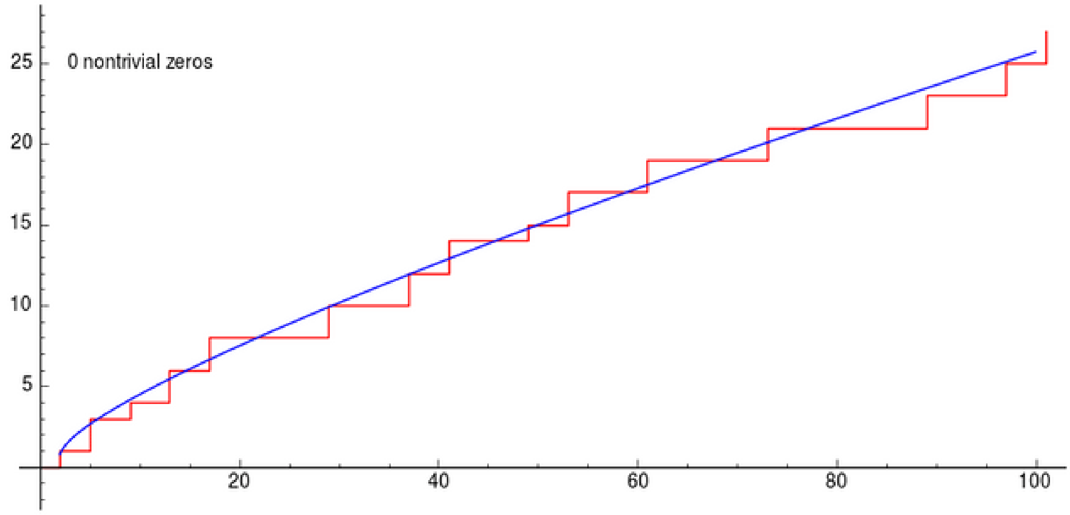
So, what about Primes?

Okay, so I'm not going to delve deeper as this article is far far longer than I thought it would be. But here's a very brief overview.

This is the prime counting function, it's a simple function that jumps up by log(p) whenever you encounter a prime, where p is the encountered prime. It's like a staircase that spikes whenever a prime number is encountered:



There is another function that exists, the Gauss function, 1/log(x) and if you chart it out, it closely follows the above prime staircase function: the blue line is the Gauss function.

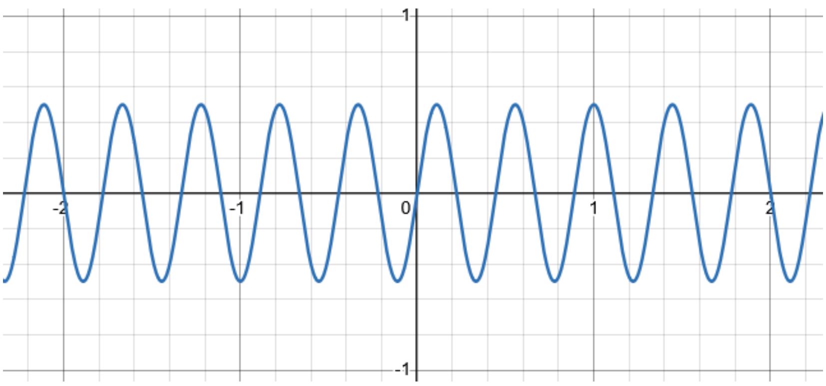


This was discovered by mathematician Friedrich Gauss. There's clearly some sort of correlation here but it's not perfect...

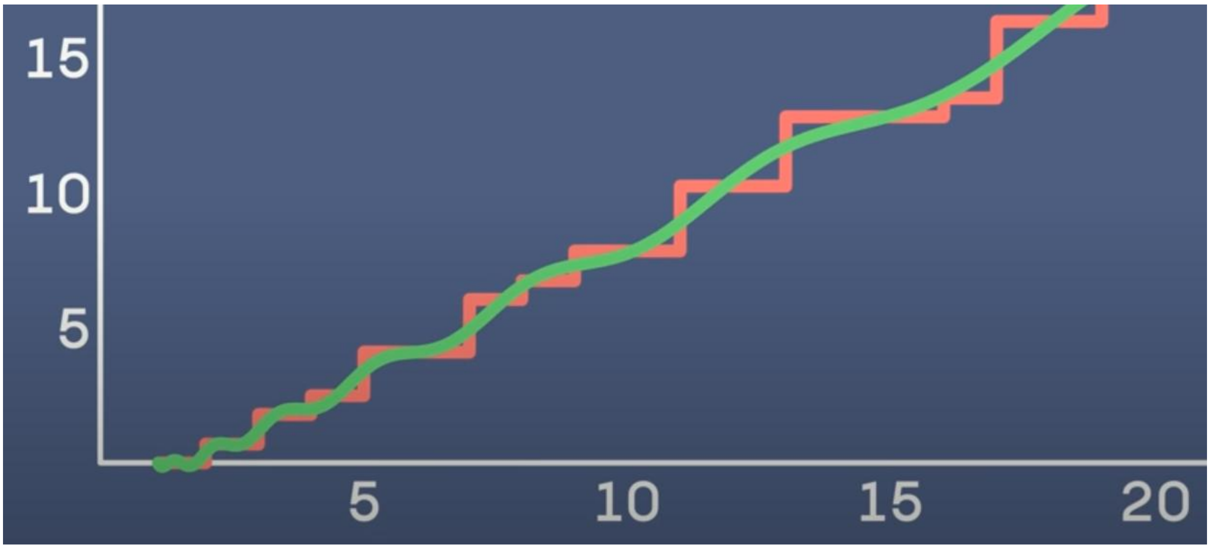
Enter Riemann.

Let's choose one of the non-trivial zeroes of the Riemann Zeta function, say: ½ + 14.13i.

Now let's create a simple waveform from this zero, where the real part of the zero is the amplitude and the imaginary part of the zero is the frequency. So ½ would be the amplitude and 14.13Hz would be the frequency. This is how it would look like:

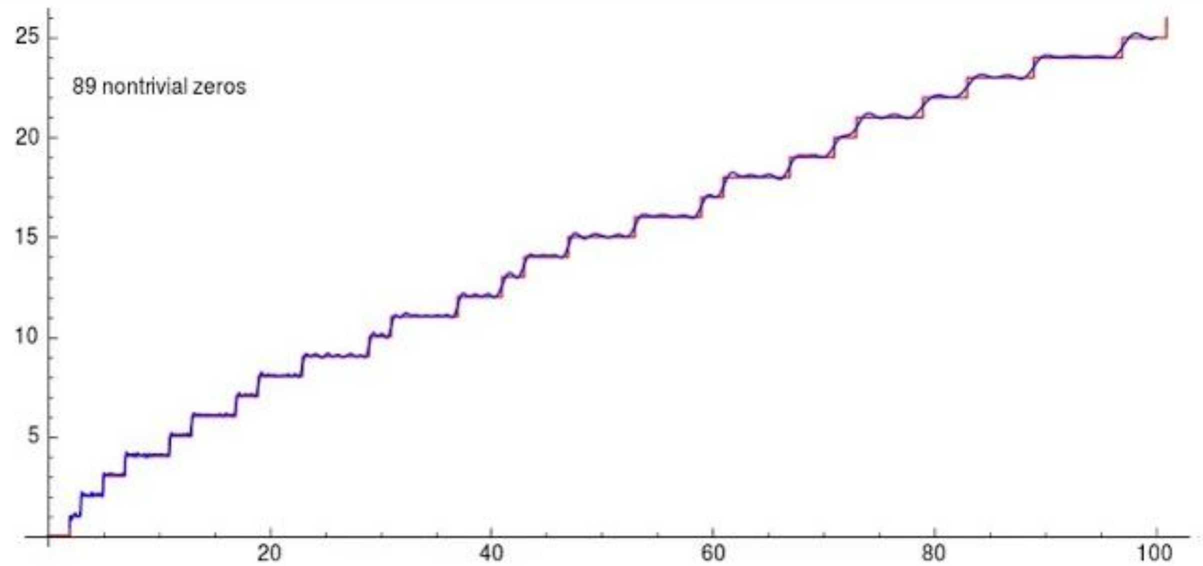


Let's superimpose this wave onto the Gauss function, 1/log(x), and check how it compares to the prime counting function:



There's a small correction here that makes the superimposed Gauss function track the prime counting function better.

Let's do the same thing for 89 non-trivial zeroes of the Riemann Zeta function and superimpose all of the waveforms onto the Gauss function:



The superimposed function becomes so much closer to the prime counting function!

A definitive equation we can use to completely predict where the next prime number is!

This is the reason why the Riemann Hypothesis is so goddamn important. If we can actually prove it, not just confirm with observations, then that would lead a number theory revolution.

There is SO MUCH I simply could not cover in this article. From the derivation of the solution of the zeta function, to the math behind the Analytical Continuation, to the Gauss Convergence of Primes...a lot has been omitted and skipped over. You can look into this on your own and I encourage you to do so after reading this article. My hope is that I spark a curiosity within and serve as a small introduction to the behemoth that is the Riemann Hypothesis.

If you have taken the time and were patient enough to suffer through this, thank you.

Have a great month ahead and goodbye till the next we meet.