Problems with the "Fundamental Theorem of Algebra".

When students learn about the Fundamental Theorem of Algebra they clearly understand the basic idea but they soon get the feeling that the theorem only works in special cases where the roots are real numbers.

This article is written for educators who wish to use the following innovative method to present this theorem from a completely new perspective.

It is important that articles of this type should be written in "*plain language*" so that they can be easily understood by people at varying levels of mathematical knowledge.

Basically, the theorem states that equations like:

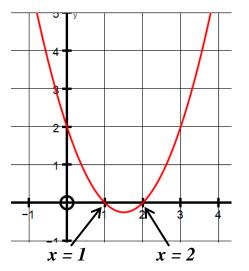
- 1. $x^2 3x + 2 = 0$ will always have 2 solutions
- 2. $x^3 7x^2 + 14x 8 = 0$ will always have 3 solutions
- 3. $x^4 5x^2 + 4 = 0$ will always have 4 solutions
- 4. $x^7 + ax^6 + bx^5 \dots + c = 0$ will always have 7 solutions
- 5. $ax^n + bx^{n-1} + cx^{n-2} + \dots + p = 0$ will always have *n* solutions

The above equations are called **polynomial equations** (powers of x) and, simply stated, the theorem states that if x^n is the highest power, the equation will have n solutions (where n is a positive integer).

The theorem needs to be closely related to graphs to get a clear understanding.

Here are three straightforward cases where the theorem applies nicely without any ambiguity.

1. Consider the graph of $y = x^2 - 3x + 2$ which is a simple parabola. The question: "When is $x^2 - 3x + 2 = 0$?" is like saying: "When is y = 0?" which clearly means "When does the graph cross the x axis?"



This is the graph of $y = x^2 - 3x + 2$ which clearly crosses the *x* axis at x = 1 and at x = 2

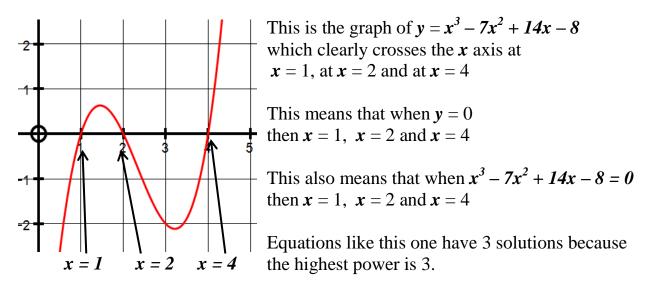
This means that when y = 0then x = 1 and x = 2

This also means that when $x^2 - 3x + 2 = 0$ then x = 1 and x = 2

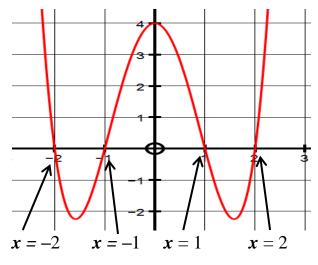
Equations like this one have 2 solutions because the highest power is 2.

2. Now consider the graph of $y = x^3 - 7x^2 + 14x - 8$ which is called a "cubic" curve.

The question: "When is $x^3 - 7x^2 + 14x - 8 = 0$?" is like saying: "When is y = 0?" which means "When does the graph cross the x axis?"



3. Now consider the graph of $y = x^4 - 5x^2 + 4$ which is called a quartic curve. As stated before, the question: "When is $x^4 - 5x^2 + 4 = 0$?" is like saying: "When is y = 0" which clearly means "When does the graph cross the x axis?"



This is the graph of $y = x^4 - 5x^2 + 4$ which clearly crosses the x axis at x = -2, x = -1, x = 1 and x = 2

This means that when y = 0then x = -2, x = -1, x = 1 and x = 2

This also means that when $x^4 - 5x^2 + 4 = 0$ then x = -2, x = -1, x = 1 and x = 2

Equations like this one have 4 solutions because the highest power is 4.

At this point of course we expect such questions as "But what if the graph does not cross the x axis? The theorem does not seem to work!"

The usual "answer" to this type of question is that "*The solutions are imaginary numbers but the equation still has the right number of solutions*"!

Another common question concerns the case when the graph just **touches** the *x* axis and students are merely told "*This is called a double solution!*" These sort of inadequate, unconvincing answers make students lose any respect for the theorem but there is a very nice way to show that the graphs actually do have <u>**physical places**</u> where the *y* value is zero!

Fig 1 This is the basic graph of $y = x^2$ and if only **real values** of *x* are used then only positive values of y are obtained. Fig 1 $x = \pm 1$ we get y = 1 $x = \pm 2$ we get y = 4 $x = \pm 3$ we get y = 9However, allowing values of *x* such as: $x = \pm i$ then y = -1and these are $x = \pm 2i$ then y = -4real y values! $x = \pm 3i$ then y = -9Fig 2 The insight, is to allow a Complex x PLANE but with just a Real y AXIS. Fig 2 Complex x plane This produces a sort of "phantom" parabola consisting of a underneath the basic parabola and at right Real x axis angles to it. and an Imaginary Nearly ALL curves have these extra $\cdot x$ axis "phantom" parts and more importantly, this has an intriguing connection with the Fundamental Theorem of Algebra.

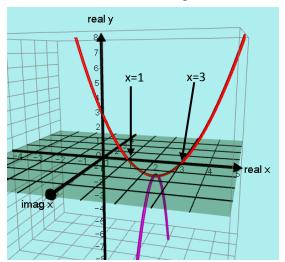
Basically, the **Fundamental Theorem of Algebra** states that polynomial equations of the form: $ax^n + bx^{n-1} + cx^{n-2} + \dots px^2 + qx + r = 0$ will have *n* solutions. (where *n* is a positive integer). This is often interpreted as:

"The solutions of an equation f(x) = 0 are where the graph of y = f(x) crosses the x axis". But this only finds the solutions which are REAL numbers! Fig 3

Consider the equation $x^2 - 4x + 3 = 0$ The graph of $y = x^2 - 4x + 3$ is as shown in **Fig 3**

The graph crosses the x axis at x = 1 and x = 3 so the solutions are x = 1 and 3

In this case, the phantom hanging below had no part to play in this logic.



However, consider the equation $x^2 - 4x + 4 = 0$ The graph of $y = x^2 - 4x + 4$ is as shown in Fig 4.

In this case, the top half of the parabola crosses the *x* axis at x = 2 **AND** the bottom half of the parabola (*the phantom*) ALSO crosses the *x* axis at x = 2. (*This is called a "double solution"*)

The graph goes through the point (2, 0) <u>TWICE</u>. The red basic parabola goes through (2, 0) and the purple *phantom* parabola also goes through (2, 0)

Fig 5 Of course, the most interesting case is when the basic **top half** of the parabola would not normally cross the *x AXIS* at all but its *phantom* would cross the *complex x PLANE*!

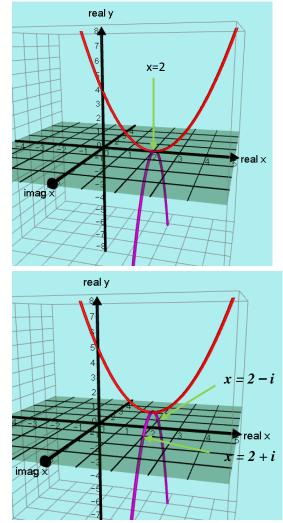
Consider the equation $x^2 - 4x + 5 = 0$ The graph of $y = x^2 - 4x + 5$ is as in Fig 5.

Clearly, any parabola of the form:

exactly two times. Fig 6.

 $y = ax^{2} + bx + c$ (with its *phantom*) will cross **any** horizontal plane y = c(which represents any real y value)

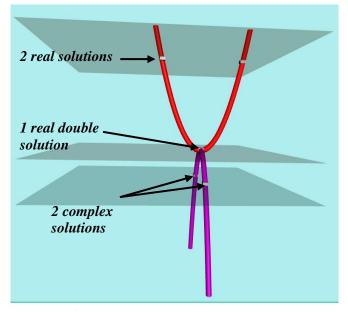
The *phantom* crosses the *x* plane at x = 2 + i and x = 2 - i as shown in Fig 5 and these are the complex solutions of the equation.



The Fundamental Theorem of Algebra can now re-stated as:

"The solutions, whether they are real or complex, of an equation f(x) = 0, are where the graph of y = f(x) crosses the complex x plane".

Fig 6.



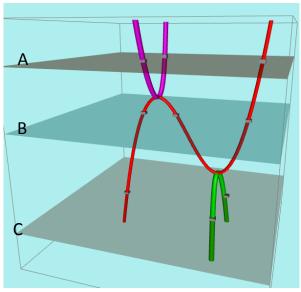
CUBIC EQUATIONS

In my website <u>www.phantomgraphs.weebly.com</u> you will find detailed working to show HOW **CUBIC** functions of the form: $y = ax^3 + bx^2 + cx + d$ each have 2 phantoms emanating from their maximum and minimum points. See Fig 7. (*The method showing how the equations of the Phantoms are calculated is explained later in this chapter.*)

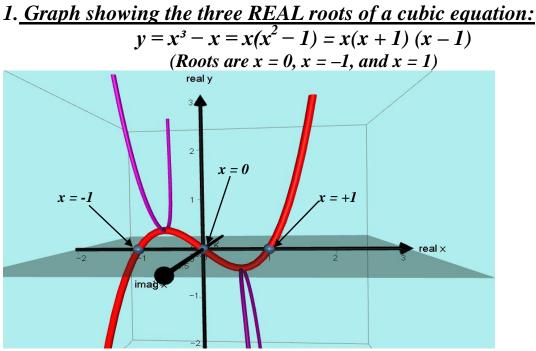
The theorem says that any Cubic equation of the form: $ax^3 + bx^2 + cx + d = 0$ will have 3 solutions.

Cubics can have **3 REAL solutions** as for the intersections with the middle **Plane B** in **Fig 7** and sometimes we have **1 real solution and 2 complex solutions** as on **Planes A and C.** See **Fig 7**.





The following is a discussion of the three basic cases for the solutions of cubic equations:



In this case the basic **RED** curve has two turning points from which the two **PURPLE** phantoms emerge.

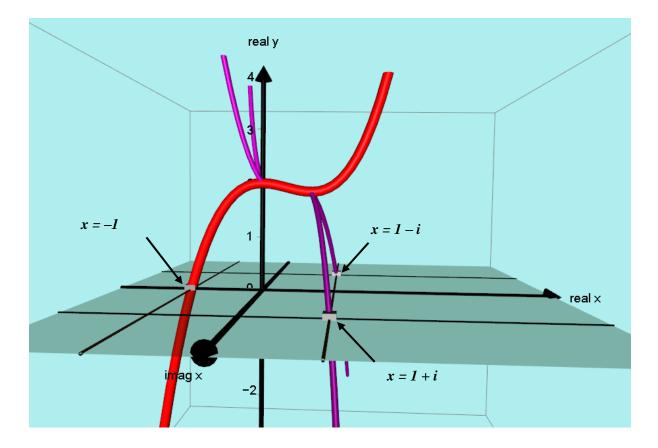
Both the upper and lower phantoms do not cross the x plane so they play no part in the solutions of $x^3 - x = 0$

The basic **RED** curve crosses the x plane THREE times showing the real solutions to $x^3 - x = 0$ are x = 0, x = -1, and x = 1

2. Graph showing one real and two complex roots of the equation:

$$y = x^{3} - x^{2} + 2 = (x + 1)(x^{2} - 2x + 2)$$

= (x + 1) (x - (1+i)) (x - (1 - i))
(Roots are x = -1, x = 1 + i and x = 1 - i)



In this case the basic **RED** curve has two turning points from which the two **PURPLE** phantoms emerge.

The upper phantom does not cross the x plane so this does not show any solutions to y = 0.

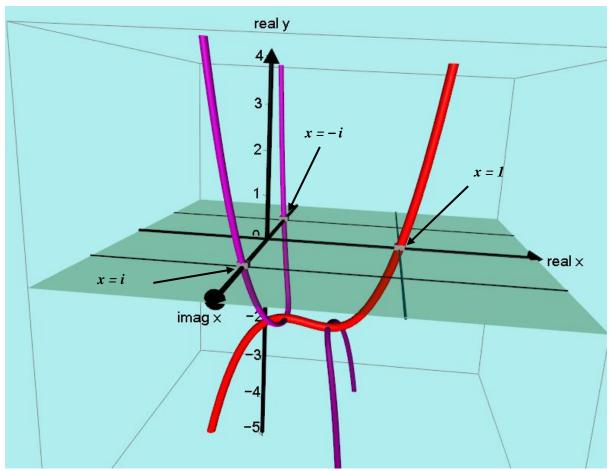
The lower phantom crosses the x plane showing the solutions x = 1 + i and x = 1 - i. Notice that because of the positions of the phantom graphs, they will always produce roots which are complex conjugates (ie a + ib and a - ib)

The basic **RED** curve crosses the x plane showing the real solution at x = -1.

3. <u>Graph showing one real and two complex roots of the equation:</u>

$$y = x^{3} - 2x^{2} + x - 2 = (x^{2} + 1)(x - 2)$$

= $(x + i)(x - i)(x - 2)$
(Roots are $x = i, x = -i$ and $x = 2$)



In this case the basic **RED** curve has two turning points from which the two **PURPLE** phantoms emerge.

The lower phantom does not cross the x plane so this does not show any solutions to y = 0.

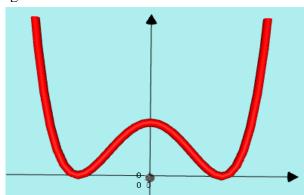
The upper phantom crosses the x plane showing the solutions at x = i and x = -i.

The basic **RED** curve crosses the x plane showing the real solution at x = 2.

QUARTIC EQUATIONS

This is a typical **QUARTIC** graph. **Fig 8** showing 1 maximum and 2 minimum points.

Fig 8.



This is the same Quartic graph with its 3 *phantoms* emanating from each turning point. **Fig 9**. The intersection points with the 4 planes A, B, C

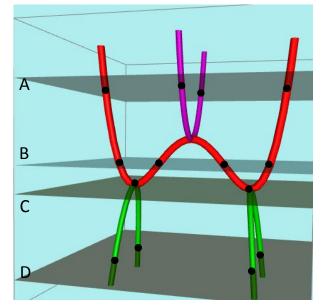
and D with the graph, are marked.

<u>Plane A</u> shows 2 real solutions on the basic RED curve and 2 imaginary solutions on the PURPLE *phantom*.

<u>Plane B</u> shows 4 real solutions on the basic RED curve.

<u>**Plane C**</u> shows 2 double real solutions which lie on the basic red curve and the GREEN *phantoms*.

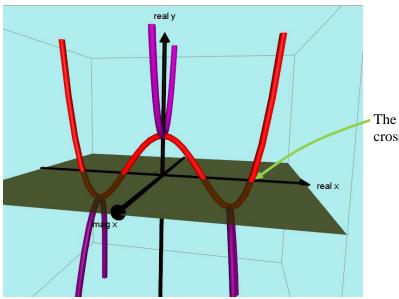
<u>**Plane D</u>** shows 4 imaginary solutions on the two GREEN *phantoms*. (2 sets of conjugate roots)</u>



Clearly a **QUARTIC** (with *phantoms*) curve will pass through ANY horizontal plane 4 times.

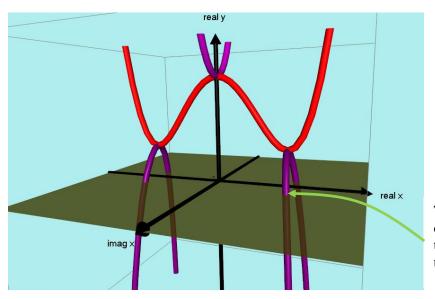
The following is a discussion of the basic cases for the solutions of quartic equations:

1. Graph showing the 4 REAL roots of a quartic equation:



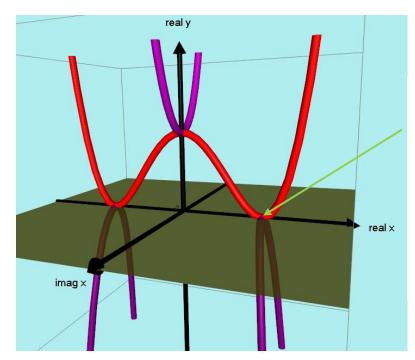
The basic red graph clearly crosses the *x* axis four times.

2. Graph showing a quartic with no real roots but 4 complex roots.



The basic red graph clearly does not cross the *x axis* but the two lower *phantoms* do cross the *complex x plane* 4 times.

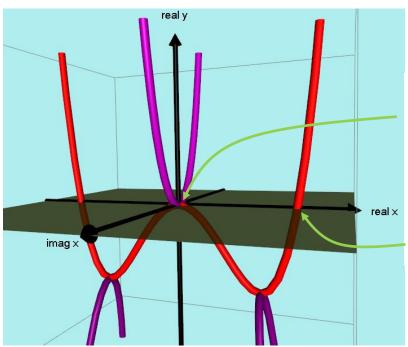
3. Graph showing a quartic with 2 sets of "double" real roots.



The basic red graph is clearly "*sitting*" on the *x* axis at 2 points but notice that the two lower phantoms also go through the same points.

The graph does cross (or touch) the x plane a total of 4 times but we sometimes say, it has 2 "double" solutions.

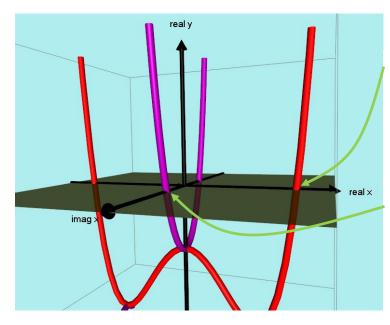
4. <u>Graph showing a quartic with 2 distinct real roots</u> and a double real root at x = 0.



The basic red graph is clearly "*sitting*" on the underside of the x axis at x = 0 and the top phantom is sitting on the top side of the x axis at x = 0 also. This is a "double" solution.

The basic red graph also crosses the x axis at two distinct points on the x axis.

5. Graph showing a quartic with 2 real roots and 2 imaginary roots.



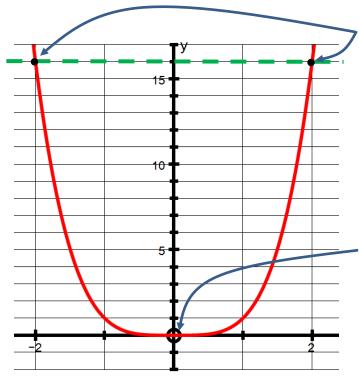
The basic red graph crosses the x axis twice which means there are 2 real roots.

The top phantom crosses the xplane (actually on the imaginary x axis) which means there are 2 imaginary roots also.

Finding the equations of the Phantom Graphs.

This would be a good stage to show the process for finding the equations of the *phantom graphs* (which use *complex x values* to still produce <u>real y values</u>).

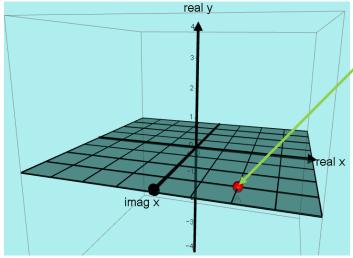
Consider the graph of $y = x^4$ As you can see this is a very U-shaped graph.



Notice that if $x^4 = 16$ then x = 2 and x = -2

But if $x^4 = 16$ then there should be 4 solutions according to the Fundamental Theorem of Algebra but only two solutions can be seen here. x = 2 and x = -2

Also, how can we have 4 solutions to the equation $x^4 = 0$? The graph only crosses this point once! It is necessary to allow some imaginary x values so an imaginary x axis is needed, as shown below.



Example: The point (2 + 3i, 0)will be at this position.

Using a 3D coordinate system the point is really (2, 3, 0)

NB

Start from the origin go 2 units along the real x axis then go 3 units along the imaginary x axis then don't move up because y = 0The point is ON the complex x plane

The **real** variable "x" must be replaced by the **complex** variable "x + iz" so the equation becomes:

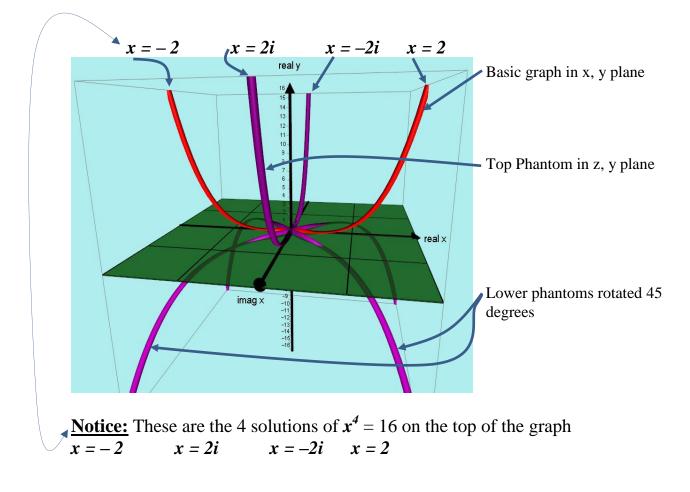
 $y = (x + iz)^{4}$ Expanding: $y = (x + iz)^{4}$ $y = x^{4} + 4x^{3}zi + 6x^{2}z^{2}i^{2} + 4x z^{3}i^{3} + i^{4}z^{4}$ Separating into real and imaginary parts this becomes: $y = x^{4} - 6x^{2}z^{2} + z^{4} + i(4x^{3}z - 4xz^{3}) - Equation A$ **The important step here is to realise that only the complex x values which produce REAL y values are required.** This means the **imaginary part of y has to be zero** in *Equation A* above! So $4x^{3}z - 4xz^{3} = 0$ $4xz(x^{2} - z^{2}) = 0$

This means z = 0 OR x = 0 OR $z^2 = x^2$

If z = 0 then substituting in Equation A this produces $y = x^4$ which is the basic equation in the x, y plane.

If x = 0 then substituting in Equation A this produces $y = z^4$ which is a phantom graph the same shape as $y = x^4$ but in the z, y plane perpendicular to $y = x^4$ which is in the x, y plane.

If $z^2 = x^2$ which means $z = \pm x$ then $y = x^4 - 6x^4 + x^4 = -4x^4$ Surprisingly, this represents 2 more phantoms the same shape as $y = x^4$ but rotated through 45 degrees.



This is the graph of $y = x^4$ for <u>real</u> y values (but allowing <u>complex</u> x values).

Recall earlier that a more meaningful way to think of the theorem is:

"The solutions, whether they are real or complex, of an equation f(x) = 0, are where the graph of y = f(x) crosses the complex x plane".

For the case of $x^4 = 0$ the graph $y = x^4$ does cross the complex x plane 4 times at x = 0. It certainly does sound silly to say the equation has 4 "solutions" but actually it has 4 "intersections" and each intersection is at x = 0.

Notice that if $x^3 = 8$ then x = 2But if $x^3 = 8$ then there should be 3 solutions according to the Fundamental Theorem of Algebra but we can only see 1 solution here at x = 2Also, how can we have 3 solutions to the equation $x^3 = 0$? The graph only goes through this point once!

Now consider the graph of $y = x^3$ and the equations of its *phantoms*.

Again replace the **real** variable x by the complex **variable** x + iz so the equation becomes:

 $y = (x + iz)^3$ because this allows for complex x values but only real y values. Expanding we get:

$$y = (x + iz)^{3}$$

$$y = x^{3} + 3x^{2}zi + 3x z^{2}i^{2} + i^{3}z^{3}$$

Separating into real and imaginary parts this becomes:

$$y = x^{3} - 3x z^{2} + i(3x^{2}z - z^{3}) - Equation A$$

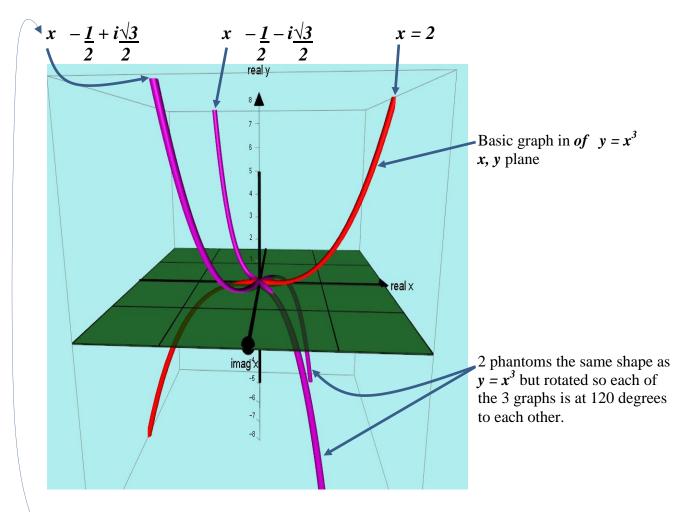
The important step here is to realise that only the complex x values which produce REAL y values are used.

This means the **imaginary part of** y has to be zero in Equation A above! So $3x^2z - z^3 = 0$

 $z(3x^2 - z^2) = 0$ This means z = 0 OR $z^2 = 3x^2$

If z = 0 then substituting in Equation A we get $y = x^3$ which is the basic equation in the x, y plane.

If $z^2 = 3x^2$ which means $z = \pm x\sqrt{3}$ then $y = x^3 - 3x3x^2 = -8x^3$ Surprisingly, this represents 2 more phantoms the same shape as $y = x^3$ but rotated through 120 degrees to each other.



This is the graph of $y = x^3$ for <u>real</u> y values (but allowing <u>complex</u> x values).

Notice: These are the 3 solutions of $x^3 = 8$ on the top of the graph: $x - \frac{1}{2} + i\frac{\sqrt{3}}{2}$ $x - \frac{1}{2} - i\frac{\sqrt{3}}{2}$ x = 2

And considering the equation $x^3 = 0$ it is clear that all 3 branches of the curve do go through x = 0.

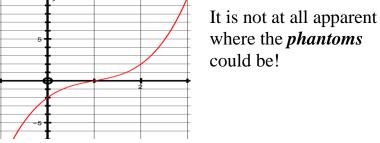
For the case of $x^3 = 0$ the graph $y = x^3$ does cross the complex x plane 3 times at x = 0.

It is nice to know that these *phantoms* seem to appear at the turning points but some curves do not have any turning points!

For instance consider a cubic with roots x = 1, 1 + i and 1 - iThe equation would be y = (x - 1)(x - (1 + i))(x - (1 - i))

be
$$y = (x - 1)(x - (1 + i))(x - (1 - i))(x$$

The basic 2D graph looks like this:

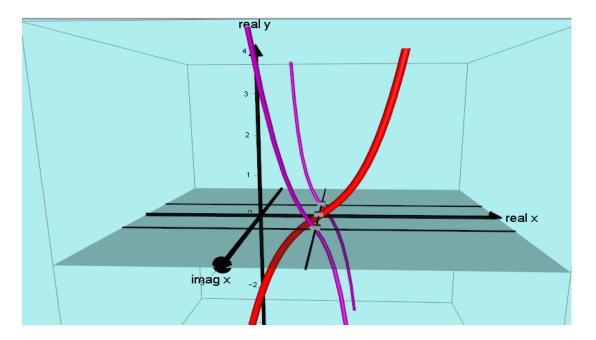


Again replace the *real variable x* with the *complex variable* (x + iz)The equation is now $y = (x + iz)^3 - 3(x + iz)^2 + 4(x + iz) - 2$ Expanding: $y = [x^3 - 3xz^2 - 3(x^2 - z^2) + 4x - 2] + i[(3x^2z - z^3) - 6xz + 4z]$

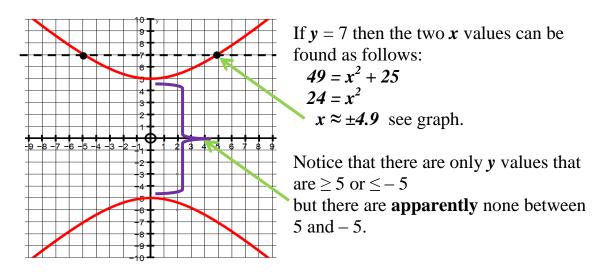
but we only want *real y values so* $3x^2z - z^3 - 6xz + 4z = 0$ factorising: $z (3x^2 - z^2 - 6x + 4) = 0$ this means z = 0 (which produces the basic 2D graph) ie $y = x^3 - 3x^2 + 4x - 2$ or $z^2 = 3x^2 - 6x + 4$ which will produce the phantom graphs by

substituting z^2 into $y = x^3 - 3xz^2 - 3(x^2 - z^2) + 4x - 2$

This is what the whole graph is like showing roots x = 1, 1 + i and 1 - i where the graphs cross the complex x plane:



It seems that when there are "missing" *y* values on a graph, there must be some *phantom graphs* which "hide" the complex solutions.

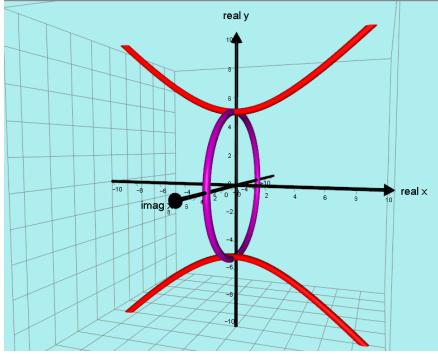


However, it is possible to find *x* values for $-5 \le x \le 5$

A very good example is the hyperbola $y^2 = x^2 + 25$

If $y = 4$ then $16 = x^2 + 25$	۰	2
$-9 = x^{2}$	$-16 = x^2$	$-25 = x^2$
$x = \pm 3i$	$x = \pm 4i$	$x = \pm 5i$

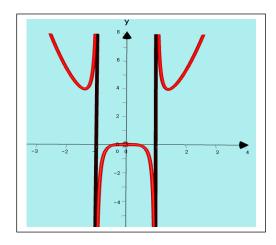
Of course, these are points on a *phantom* CIRCLE of radius 5 units, in the plane perpendicular to the hyperbola!



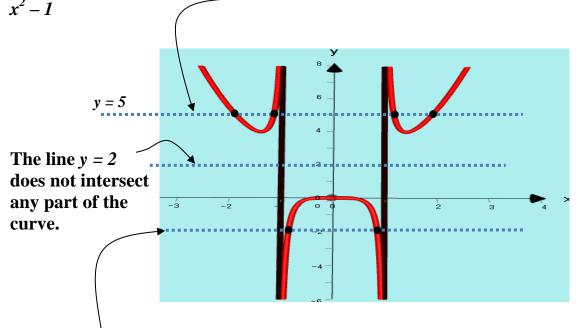
Now there are no "missing" *y* values because the graph will cross any horizontal plane exactly two times. By far the most challenging complex algebra was needed in finding the equations of the *phantoms* for the function $y = \frac{x^4}{x^2 - 1}$

Clearly there are no real values of y in the interval 0 < y < 4

Interestingly, considering a general *y* value such as y = c then $\frac{x^4}{x^2 - 1} = c$ which produces a **typical quartic** equation $x^4 - cx^2 + c = 0$ which of course has **4 solutions**.



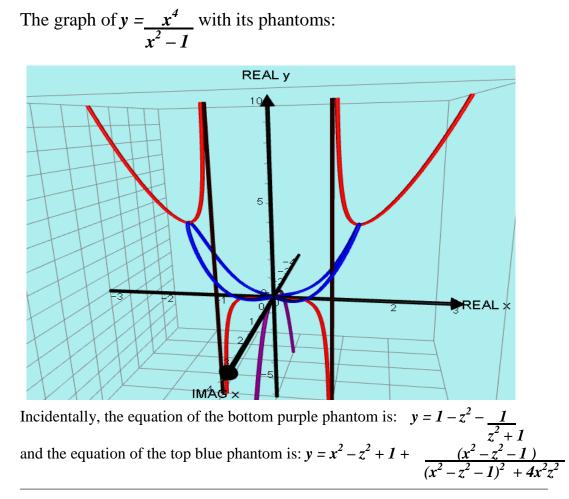
Notice that a line such as y = 5 (see below) crosses 4 times so the equation $\frac{x^4}{x^2 - 1} = 5$ clearly has 4 real solutions.



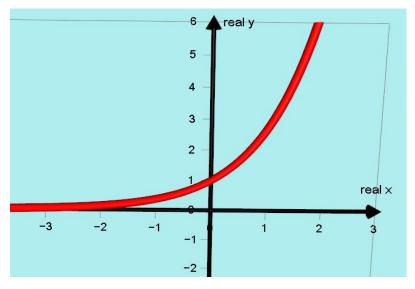
The line y = -2 crosses TWO times so now there are apparently only **two** real solutions.

But when the *Phantoms* are added (see graph below) it is clear that **any** horizontal plane y = c crosses the graph 4 times.

This further verifies the truth of the Fundamental Theorem of Algebra.



<u>The Exponential Graph $y = e^x$ </u> Another graph which has missing *y* values is $y = e^x$ the exponential graph! This idea that e^x could possibly be negative would never occur to most people! The basic 2D graph of $y = e^x$ is below:



Because e^x is always > 0 It is hard to accept that there could be *y* values under the x axis. But why not?

Obviously this could be possible if some *complex x* values can be found to produce *real* y values which are negative.

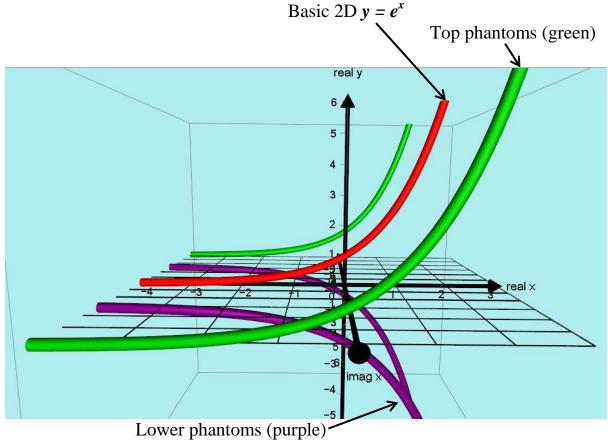
Again replacing the **real variable x** with the **complex variable x** + **iz** the equation becomes: $y = e^{x + iz} = e^x$. $e^{iz} = e^x(cosz + isinz)$ using Euler's formula.

If $\mathbf{z} = \mathbf{0}$ the equation just becomes the basic curve $y = e^x$

If $z = 2n\pi$ (i.e. any multiple of 360°) the equation becomes: $y = e^{x}(\cos 2n\pi + i\sin 2n\pi) = e^{x}(1 + \theta i) = e^{x}$ This means there are infinitely many versions of $y = e^{x}$ all parallel to the basic 2D curve spaced at intervals of 2π in the imaginary x direction.

If
$$z = (2n+1)\pi$$
 (ie multiples of $180^{\circ} + 360^{\circ}$) we get:
 $y = e^{x}(cos(2n+1)\pi + isin(2n+1)\pi)$
 $= e^{x}(-1 + 0i) = -e^{x}$

This means there are infinitely many versions of $y = e^x$ but **upside down** and all parallel to each other starting at $\pm \pi$ and then spaced at intervals of 2π to each other in the imaginary x direction.

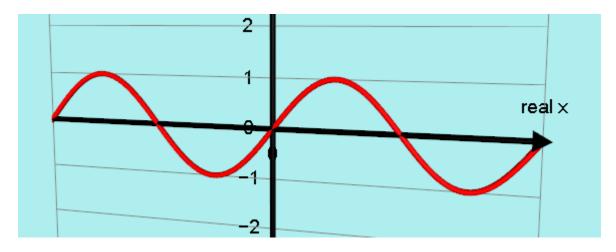


Trigonometrical Graph y = sin(x)

The basic trigonometrical graphs such as y = sin(x) and y = cos(x) only exist between 1 and -1.

It seemed nonsense that sin(x) could equal 2 or any value > 1 or < -1.

This is the basic y = sin(x) graph in 2D



Again replacing the **real variable x** with the **complex variable x + iz** the equation becomes:

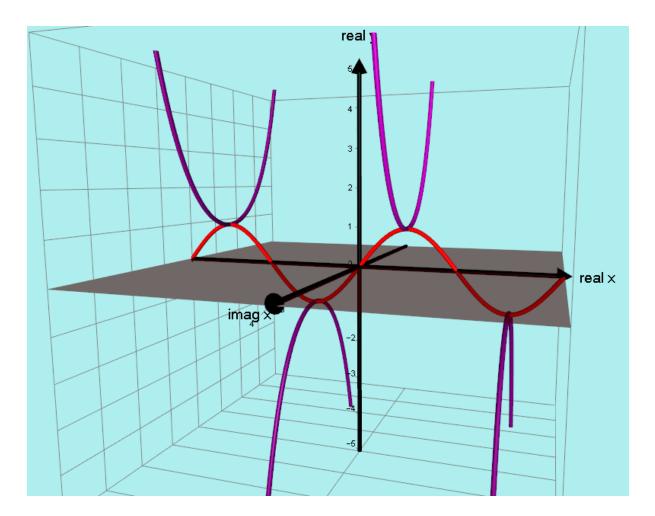
y = sin(x + iz)= sin(x) cos(iz) + cos(x) sin(iz) = sin(x) cosh(z) + cos(x) sinh(-y) = sin(x) cosh(z) - cos(x) sinh(z)

If z = 0 this just produces the basic 2D curve y = sin(x)

If $x = \frac{\pi}{2} + 2n\pi$ then $y = sin(\frac{\pi}{2} + 2n\pi + iz) = 1 \times cosh(z) - 0 \times sinh(z) = cosh(z)$ This means at every **maximum point** a *phantom* y = cosh(z) appears.

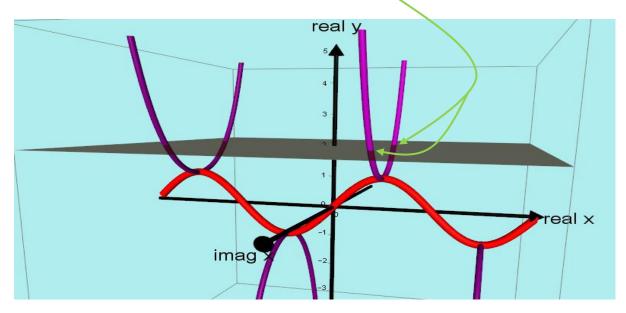
If
$$x = -\frac{\pi}{2} - 2n\pi$$
 then $y = -1 \times \cosh(z) - 0 \times \sinh(z) = -\cosh(z)$

This means at every **minimum point** a *phantom* y = -cosh(z) appears.



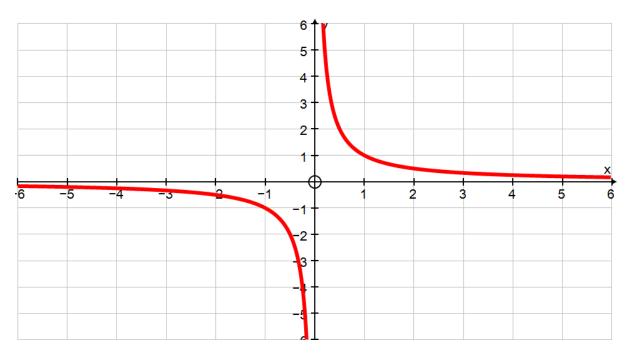
Here we can see that any equation sin(x) = c will have infinitely many solutions! We are no longer restricted to $-1 \le x \le +1$

If we restrict our domain to just one period where $0 \le x \le 2\pi$ then equations like $sin(x) = \frac{1}{2}$ will just have 2 solutions and sin(x) = 2 will also have 2 solutions.



Some graphs do not have any phantoms.

There are some graphs which do not have *phantoms*, for example: $y = \frac{1}{x}$



Notice that any horizontal line will always cross this curve **one time.** There are no "missing" values which is the usual sign that there are "*phantoms*". This is because solving $\underline{I} = c$ produces cx = 1 so $x = \underline{I}$ for which there is x c

always just one x value for all values of c (except c = 0 which is the horizontal asymptote of course).

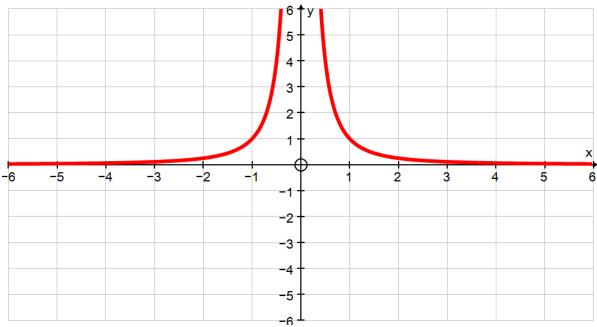
In attempting to find any phantoms by algebra, variable x is replaced by x + iz: Let $y = \frac{1}{x + iz} = \frac{1}{x + iz} \times \frac{(x - iz)}{(x - iz)} = \frac{x}{x^2 + z^2} - \frac{iz}{x^2 + z^2}$

For complex x values to produce only real y values, the imaginary part of y has to be zero.

This means $\frac{iz}{x^2 + z^2} = 0$ so z = 0 which simply means that $y = \frac{1}{x + 0i} = \frac{1}{x}$

which is just the original equation. So there are no phantoms.

However considering $y = \underline{1}$ it is obvious that there are "missing y values" below x^2



the x axis.

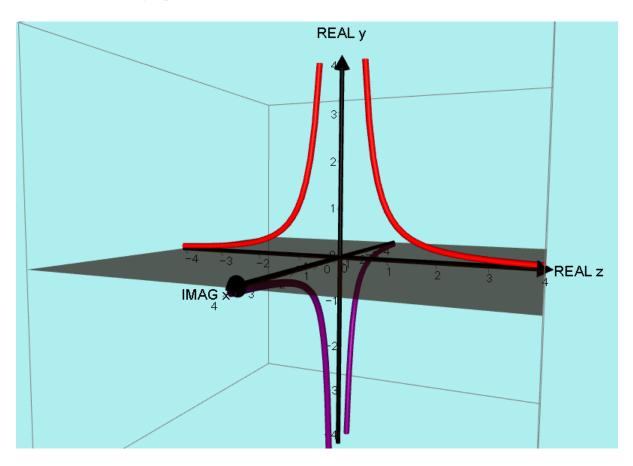
Any horizontal line y = c where c > 0 will cross this graph twice. This is because the equation $\frac{1}{r^2} = c$ becomes $x^2 = c$ which always has two solutions. To find the equation of the *phantom* graphs for $y = \frac{1}{x^2}$ the same method is used.

Let $y = \frac{1}{(x + iz)^2} = \frac{1}{(x + iz)^2} \times \frac{(x - iz)^2}{(x - iz)^2}$ $y = \frac{x^2 - z^2 - 2xzi}{(x^2 + z^2)(x^2 + z^2)}$ So $y = \frac{x^2 - z^2}{(x^2 + z^2)^2} - \frac{2xzi}{(x^2 + z^2)^2}$

For complex x values to produce only real y values, the imaginary part of y has to be zero.

This means $\frac{2xzi}{(x^2 + z^2)} = 0$

So that either z = 0 which just implies $y = \frac{1}{x^2}$ which is the original equation or x = 0 which means $y = \frac{0 - z^2}{(0^2 + z^2)^2} = \frac{-1}{z^2}$ which produces the purple phantoms as shown on the graph below:



Now any horizontal plane y = c (except c = 0) will cross this graph twice.

NB

To draw a graph with *real values of x and y* then **2 axes** are required and the graph is in **<u>2 Dimensions</u>**.

To draw a graph where y is real but x can have a real part and an imaginary part (like 3 + 4i) then 3 axes are required, namely a real y axis, a real x axis and an imaginary x axis. This requires <u>3 Dimensions</u>.

To draw a graph where *x* and *y* can both have a real part and an imaginary part we would need 4 axes. This requires <u>4 Dimensions</u> which cannot be drawn in our 3 dimensional world!