AC 2009-700: VISUAL ANALYSIS AND THE COMPOSITION OF FUNCTIONS

Andrew Grossfield, Vaughn College of Aeronautics

Throughout his career Dr. Grossfield has combined an interest in engineering design and mathematics. He studied Electrical Engineering at the City College of New York, graduating with a BSEE. During the sixties, he attended the NYU Courant Institute at night, obtaining an M.S. degree in mathematics, while designing circuitry full time during the day for aerospace/avionics companies. He earned his doctorate studying Continuum Mechanics under the direction of L. M. Milne-Thomson, CBE at The University of Arizona. He is a member of ASEE, IEEE, and MAA. Grossfield@IEEE.org is his email address.

Visual Analysis and the Composition of Functions

Andrew Grossfield
Vaughn College of Aeronautics

Abstract

A major problem today concerns educating the next generation of engineers, mathematicians and researchers. Too many of our nation's students end up neither comprehending nor liking math courses. More intensive drilling of material as currently practiced may be both ineffective and undesirable. In fact, this rigorous drilling may turn more young students away from mathematics and the sciences. Why has this situation developed when mathematics is so interesting?

There are situations and mathematical principles that will enable graphs of functions to be easily produced. This paper will provide and discuss principles that can be applied in graphing a large class of functions. The graphs of polynomials and special basic functions formed by functional composition acting on polynomials will be provided as examples of visual thinking. Engineering students who are encouraged to develop the skills of visual thinking in mathematics may find these skills beneficial in their analytical engineering studies.

A student could find pleasure and confidence in discovering the ability to gain insight into the graphical behavior of a large class of functions. This same student may become more open to studying other aspects of polynomials and other functions. These techniques can provide a quick check of computer-generated graphs or be employed when a computer is unavailable or inconvenient. If we desire to recruit more students into the analytical and other sciences, we need to discover better, easier and more pleasurable ways to present conventional math concepts before attempting to accelerate curricula by moving advanced differential concepts into the lower grades.

Contents

Negative whole number powers, reciprocals and rational functions

Fractional powers, square roots, cube roots and algebraic functions

The square of a square root and the square root of a square

Absolute values

7. Graphical oddities

   Intersections of curves

   Unions of curves

   Equations of regions

8. Classic examples of visualizations and Euler's constant
9. Conclusion

All the operations described in the paper can be verified easily by using a graphing utility. The word curve will be used to mean the graphs of piece-wise differentiable functions including straight lines and also finitely multi-valued functions.

1. Introduction

In engineering colleges during the 1950's, a student had to become acquainted with all kinds of visual constructs that were needed to solve problems of design. Oscilloscopes displayed voltage time signals; spectrum analyzers displayed signal Fourier components and curve tracers displayed diode and transistor characteristics. In addition, students contemplated such wonderful mathematical inventions as shear-moment diagrams for structural beam analysis, Mollier diagrams of steam tables, root locus and Nyquist diagrams for feedback control systems. It seemed engineers visualized everything. Graphs were mandated in engineering design and problem solving.

During the 1960's, as a young engineer in graduate school, I saw many teaching mathematicians distrusted and disparaged graphs and visual techniques. Their conventional view was that mathematical theorems could not be proved with pictures and so visualizations were dangerous. The fears of mathematicians, the constructions of Cantor and Dedekind, were rarely confronted by engineering students who needed to become more familiar with continuous and smooth functions. I tried to rely on Russian books[1], which were more likely to contain pictures and be descriptive, to gain mathematical insight. I do not believe that my fellow math majors confronted the questions that were in my mind as an engineer. By the 1990's more teaching mathematicians were willing to promote graphics and the calculus reform movement came into fashion with slogans such as, "lean and lively" and "pump not a filter."

Perhaps the rigor that was conventionally required of mathematics majors is not appropriate for engineering students. I am suggesting that math teachers consider replacing class time spent on delta-epsilon arguments with visual thinking.

This paper will examine visualization in the graphing of functions. Widely available software will help students acquire the graph of a function. Many texts have "libraries of functions" where a student can observe and memorize the graphs of functions. The graphs that are shown here are not for memorization, but to demonstrate that the properties of the curves emanate from the properties of numbers and the mechanisms of construction of the functions.

As an example, in these libraries the graph of the function, $y = \frac{1}{x}$ is displayed as a hyperbola with a pole at $x = 0$ and a horizontal asymptote. A property of division is that reciprocals of larger positive numbers are smaller and therefore reciprocals of increasing functions will be decreasing. The reciprocal of any function, which increases without bound as x increases without bound, will decrease asymptotically to zero. The pole results from the division by zero and the reciprocal of the large values of x must level off at $y = 0$, producing the horizontal asymptote. There is no alternative to the fact that the graph of the reciprocal of any increasing linear function, including $y = x$, must appear as a hyperbola.

Polynomial functions produce nice curves; curves that are defined everywhere, continuous, smooth and limited in their ability to "wiggle" by the degree of the polynomial. Allow for reciprocation and the resulting curves can acquire new features. Rational functions, obtained by dividing polynomials, can have horizontal asymptotes, point gaps and infinite jumps called poles. Computing problems occur only at a finite number of isolated points that are the zeros of the denominator. The class of functions that allow for square and cube roots are called algebraic functions. The new features of algebraic functions are:

1) intervals, called excluded, where the function is not defined and where the curve does not exist.
2) intervals where the function has more than one value or the curve doubles back on itself.
3) points with vertical tangent lines, infinite slopes.
4) The curve can intersect itself.
5) The curve can have points exhibiting a sharp change and even reversal in direction. These jumps in slope are called corners or cusps.
6) Algebraic curves, such as the ellipse, can be limited in extent, both horizontally and vertically.
7) The curve may include or consist of isolated points.

Simple examples of algebraic curves are circles, hyperbolas and tilted parabolas. Vertical parabolas are included in the category of the elementary second-degree polynomials.

Descartes invented a technique to study equations in two variables by examining curves that would illustrate the relationship between the variables. The central idea of this paper is to describe the principles used in creating, from the equation relating two variables, the invaluable mental images used in constructing, seemingly out of nothing, the sketch of the curve. A smooth (continuous and without corners) non-constant curve with two separate zeros must have a least one extreme point between them. If the zeros are double zeros then between each of the zeros and the extreme point there will be at least one point of inflection. The pictures dictate it. The knowledge of the appearance of the curve and its properties should enhance a student's confidence in studying equations in two variables.

2. Polynomials

A completely factored polynomial is easy to graph; but factoring a polynomial is not easy. To graph the polynomial we need to separate the behavior of the polynomial far from the origin from the behavior in the neighborhood of the oscillations and zeros. The regions far from the origin, where x is large and either positive or negative, can be called the remote regions. The interval between the remote regions, which contains the zeros and oscillations, can be called the near region. To graph the polynomial, the following properties of all polynomials need to be recognized:

1) Polynomials are defined at every point, are single-valued, continuous and smooth.

2) In the remote regions the behavior of the polynomial is completely determined by the leading term. When the values of x are large, the value of the leading term, $a_n x^n$, far exceeds the sum of the values of all the other terms. And therefore in the remote regions, the polynomial behaves like the leading term, $a_n x^n$, and so approaches either $+\infty$ or $-\infty$ depending on the parity of the exponent and the sign of the coefficient of the leading term.

3) A polynomial of degree n can cross a straight line no more than n times. This means that a polynomial can have no more than n zeros. This fact places an upper bound on the amount of "wiggling" possible by a polynomial of fixed degree.

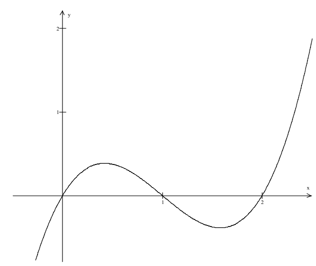


Figure 1 $y = x(x - 1)(x - 2)$

If a polynomial of degree n can be factored completely, it can be graphed by examining the sign of the leading term to find the signs of the polynomial in the remote regions. So the graph of the polynomial starts on the left at either $+$ or $-$ infinity goes continuously and smoothly through the zero corresponding to each factor and ends on the right at either $+$ or $-\infty$. As an example, consider the third degree polynomial, $y = x(x - 1)(x - 2)$, which has zeros at $x = 0$, $x = 1$ and $x = 2$. Note that when x is negative and large so is y. We see that when x is positive and large so is y. The graph must start at the lower left at negative infinity, go through the origin, cut the x-axis at $x= 1$ and then again at $x =2$ and then continue smoothly to positive infinity at the upper right. The graph is shown in figure 1.

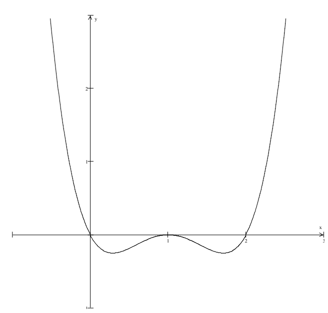


Figure 2
$y = x(x - 1)^2(x - 2)$

If a polynomial has a double zero, say $y = x(x - 1)^2(x - 2)$, then near $x = 1$ the polynomial behaves as either $\pm (x - 1)^2$ depending on the sign of the product of all of the other factors. This means that the polynomial does not cross the x axis at $x = 1$ but touches it tangentially and returns in the direction from which it came. Locally the polynomial behaves like a parabola at the zero, being either a minimum or a maximum. The values of y approach plus infinity for both very large negative and very large positive values of x. The graph of this fourth degree polynomial is shown in figure 2.

If a polynomial has a triple zero, say $y = x(x - 1)^3(x - 2)$, then near $x = 1$ the polynomial behaves as either $\pm (x - 1)^3$ again depending on the sign of the product of all of the other factors. This means that the polynomial crosses the x-axis tangentially at $x = 1$. Locally the polynomial behaves like a cubic at the zero having a point of inflection with zero slope. The zero slope is shown in the graph of this fifth degree polynomial which is plotted in figure 3.

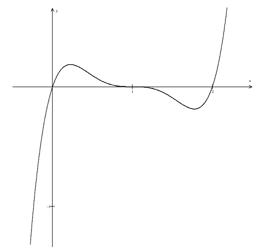


Figure 3
$y = x(x - 1)^3(x - 2)$

3. Arithmetic operations on functions and their effect on curves

Addition and Subtraction: Typical of the principles useful in curve sketching related to adding and subtracting functions are:

- Any value plus or minus zero is left unchanged.
- Adding positive values produces larger positive values.
- Large positive values plus small positive values produce slightly larger values.
- More information is needed before the results can be determined of subtracting large positive values from large positive values.
- Large values minus small positive values produce diminished large values.
- The slope of a sum (difference) of functions equals the sum (difference) of the slopes of the individual functions.
- Sums, differences and products of continuous functions are continuous. Discontinuities in quotients of continuous functions can occur only at the zeros of the denominator.
- Sums, differences, products of differentiable (smooth) functions are differentiable (smooth). If the denominator of the quotient of smooth functions is not zero, then the slopes of the quotient will be well defined, that is the quotient function will not have any cusps.
- Sums of monotonically increasing functions are monotonically increasing.

As an example add the straight line, $y_1 = x$ to the parabola, $y_2 = x^2$ to get another parabola $y_3(x) = x + x^2$. Since both functions, y_1 and y_2, are continuous and smooth, the sum y_3 will be continuous and smooth. For positive x, the positive values for y add to produce values larger than x or x^2. The parabola $y_2 = x^2$ is forced upward. For negative values of x, the parabola $y_2 = x^2$ is pulled downward. The y-intercept remains at zero. In the negative region as x moves negatively, ultimately x^2 becomes bigger than x, and the parabola must cross the x-axis and rise to the left monotonically. These graphs are shown in figure 4 and figure 5.

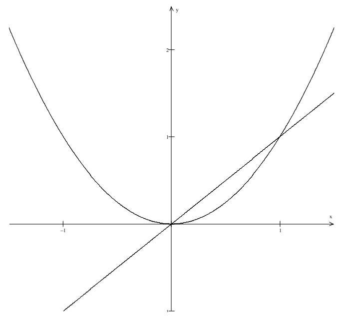


Figure 4 The graphs of $y_1 = x$ and the parabola, $y_2 = x^2$

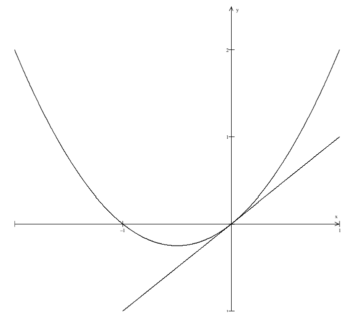


Figure 5 The graph of $y_1 = x$ and the sum of y_1 and y_2, $y_3 = x + x^2$

As a second example add $y_1 = x$ to $y_2 = 0.5 \sin(5x)$ to get $y_3(x) = x + 0.5 \sin(5x)$. y_1 produces an upward drift while $y_2 = 0.5 \sin(5x)$ is an oscillation. The result, y_3, is an oscillation that drifts upward. At the zeros of the sinusoid, the straight line, $y_1 = x$, intersects the sum, $y_3(x)$.

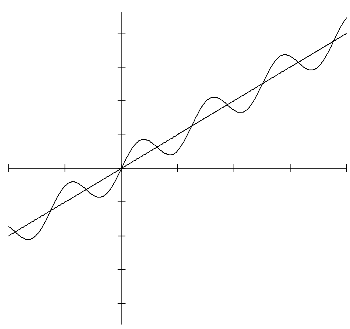


Figure 6 $y_3(x) = x + 0.5 \sin(5x)$

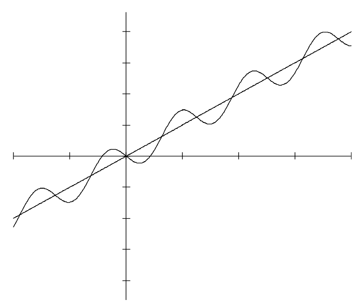


Figure 7 $y_4(x) = x - 0.5 \sin(5x)$

To subtract functions, flip the minuend (that is, the function being subtracted) and add. The graph of $y_4(x) = x - 0.5 \sin(5x)$ is displayed in figure 7.

Multiplication: Typical of the principles useful in sketching related to multiplying functions are:

- The product of finite values and zero equals zero.
- The product of functions without poles possesses all the zeroes of each of the functions.
- The product of two functions of like sign is positive.
- The product of two functions of differing sign is negative.
- The product of even symmetric functions has even symmetry.
- The product of two odd symmetric functions has even symmetry.
- The product of an even symmetric function and an odd symmetric function has odd symmetry.

In figure 8 three straight lines are graphed with zeros at x = 0, x =1 and x = 2.

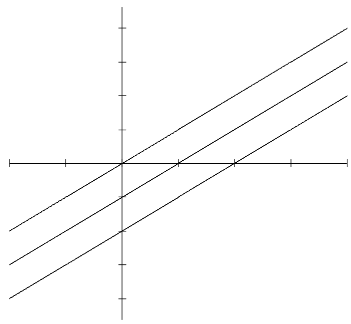


Figure 8 y = x, y = x − 1 and y = x − 2

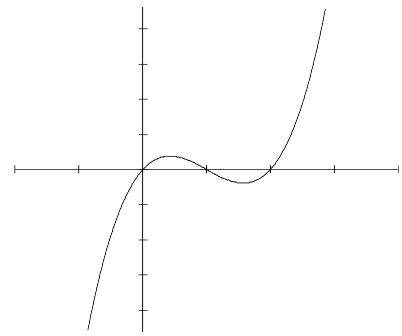


Figure 9 y = x(x−1)(x−2)

Their product y = x(x−1)(x−2), is shown in figure 9. In accordance with the principles, their product must be continuous and smooth. Since zero times any finite number is zero, the product must be zero whenever any of the factors is zero.

For x > 2	increasing x increases the product.
For 1 < x < 2,	only one factor is negative and therefore the product must be negative
For 0 < x < 1,	there are two negative factors and therefore the product must be positive
For x < 0,	the product of three negative factors is negative.

Division: Typical of the principles useful in obtaining a sketch of the quotient of functions are:

- Dividing finite values by larger values produces smaller values.
- Dividing finite values by smaller values produces larger values.
- Dividing non-zero values by zero produces a pole.
- The principles regarding signs for multiplication hold for division.
- The symmetry principles of multiplication hold for division.

4. Elementary operations on curves and the algebra needed to produce them

Translations: Translating a curve means moving the curve, horizontally, vertically or both without rotating or altering its shape. Translating is easy to visualize and also easy to accomplish in algebraic form. If the curve is expressed explicitly, that is, in the form y = f(x), to raise it vertically, add a positive constant to the expression for y. To lower the curve vertically, subtract a positive constant from the expression for y. To move the curve horizontally h units to the right, substitute x − h wherever x appears in the expression for f(x).

To translate a curve expressed implicitly as F(x, y) = 0, both k units upward and h units to the right, substitute $x - h$ for x and $y - k$ for y. As an example, the circle of radius 2, centered at the origin is described algebraically by the equation: $x^2 + y^2 = 4$.
To translate the circle 3 units upward and 5 units to the left, change the equation to: $(x + 5)^2 + (y - 3)^2 = 4$.

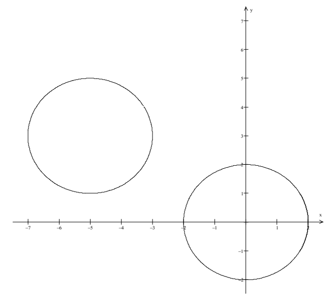


Figure 10

A circle at the origin and translated

Stretches, compressions and flips

If a curve is expressed explicitly, that is in a form y = f(x), the curve can be uniformly stretched by multiplying all the y values by some number say r > 1. So we can stretch the parabola $y = x^2$ 5 times by writing $y = 5x^2$. The circle of radius 2 can be stretched 3 times by modifying its double–valued equation $y = \pm(4 - x^2)^{1/2}$ to $y = \pm 3\,(4 - x^2)^{1/2}$. The curve is now an ellipse with a vertical major semi-diameter of 6 . The horizontal intercepts remain unchanged at ± 2. If the multiplier is between 0 and 1, then the curve will be uniformly compressed vertically. The graphs for the vertically stretched and compressed circle are shown in figures 11 and 12.

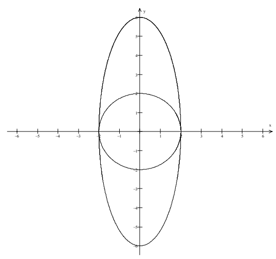


Figure 11
A circle of radius two
and after stretching vertically
by a factor of 3

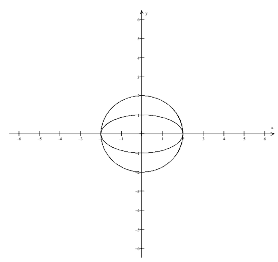


Figure 12
A circle of radius two
and after compressing vertically
by a factor of ½

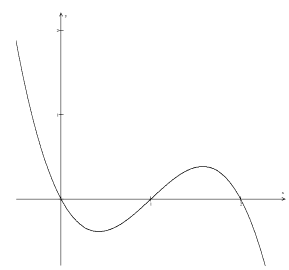


Figure 13
$y = -x(x - 1)(x - 2)$
Flipping a cubic

If the multiplier is −1, then the curve will be flipped vertically. The graph of the cubic polynomial $y = x(x - 1)(x - 2)$ when flipped vertically is shown in figure 13. After flipping, vertical values that were positive are now negative and vice versa.

A curve can be uniformly compressed horizontally by replacing every x in the equation with the value ax where a > 1. Electrical engineers write the equation $v(t) = \sin(\omega t)$ to describe a sinusoidal voltage signal varying with time t, and a radian frequency $\omega = 2\pi f$ where f represents cycles per second. Increasing ω means more cycles per second, which produces shorter periods, that is, the sinusoid is compressed on the time axis. The graphs of $y = \sin(t)$ and $y = \sin(3t)$ are displayed in figures 14 and 15.

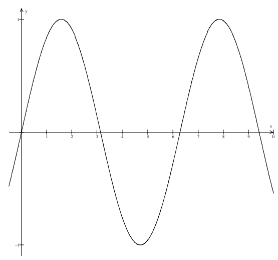


Figure 14
v(t) = sin (t)

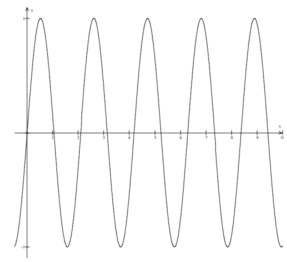


Figure 15
v(t) = sin(3t)

5. Chains, the composition of functions

Consider the concatenation of functions: w = f(y); y = g(x) or w = f(g(x)). Is the variable w, dependent on the variable x? The answer is yes. Since the variable x controls the variable y and the variable y controls the variable w, ultimately x controls w. Usually varying x will cause w to vary. There can be graphs for y = g(x), w = f(y), and w(x). We will call f(y) the outer function, g(x) the inner function and w = f(g(x)), the composition or chain form. Now we are in a position to explore the visual composition of functions.

It should be understood that the operation of chaining affects the vertical or y values of the curve of the inside function. Suppose instead of the explicit forms, w = f(y); y = g(x), the inside function is given implicitly by F(x, y) = 0 , the implicit form of the chain will be:
$F(x, f^{-1}(w)) = 0$.

In the remainder of this paper, the effect on polynomials of composition with various simple outer functions will be considered. If the outer functions are applied to periodic functions, the resulting composite functions will be periodic. The case where the outer function is linear w = my + b has already been considered. In this case the curve for y = g(x) will be vertically stretched, compressed and/or perhaps flipped depending on m and vertically translated by the value of b. The remaining outer functions that we will consider are:

1. squaring $w = \{g(x)\}^2$
2. cubing $w = \{g(x)\}^3$
3. taking the reciprocal $w = \dfrac{1}{g(x)} = \{g(x)\}^{-1}$
4. taking the square root $w = \sqrt{g(x)} = \{g(x)\}^{1/2}$
5. taking the cube root $w = \sqrt[3]{g(x)} = \{g(x)\}^{1/3}$
6. taking the absolute value $y = |g(x)|$
7. the cases where w = f(y) and y = g(x) exhibit even or odd symmetries

Principles of Chains or Composition

There are properties of functions that are preserved under composition. Students should be acquainted with the following principles:

- Chains of continuous functions are continuous.
- Chains of differentiable functions are differentiable.
- Monotonic functions of monotonic functions are monotonic.
- General functions of periodic functions are periodic.
- Periodic functions of general functions are not periodic, e.g. $y = \sin(x^2)$ is not periodic.
- Periodic functions of linear functions are periodic.
- The composite of an even symmetric function with a function possessing half-wave symmetry will have a period equal to half the period of the half-wave symmetric function.
- Even symmetric functions of either even or odd functions exhibit even symmetry.
- Odd symmetric functions of even functions exhibit even symmetry.
- Odd symmetric functions of odd functions exhibit odd symmetry.

If $y = g(x)$, the derivative of the chain df/dx of f(g(x)) equals the product, when both are defined, of the derivatives of f(y) and g(x); that is, $\frac{df}{dx} = \frac{df}{dy}\frac{dy}{dx}$. From this it follows that if either f(y) or g(x) has a zero slope then so does the composition f(g(x)).

6. Functions of polynomials

Squaring

The principles relating to squaring a function are:

- The square of any function is non-negative.
- The square of a differentiable or continuous function will be differentiable or continuous.
- If $x = 0$, then $x^2 = 0$ and if $|x| = 1$, then $x^2 = 1$.
- Small numbers squared become smaller still; that is, if $0 < |x| < 1$, then $x^2 < |x|$.
- Large numbers squared become still larger; that is if $1 < |x|$ then $|x| < x^2$.
- If a continuous function has a positive maximum or a negative minimum then the square will have a maximum at the same horizontal value.
- A tilted linear function squared becomes parabolic. Squaring a function with a linear zero produces a function that has a horizontal tangent at the same horizontal value.

To see why the square of a tilted line appears parabolic, just check:

At the horizontal intercept of the straight line, the square must be zero.

Near the horizontal intercept, on either side, the values of the line are small so therefore the values of the square must be even smaller.

Farther away from the intercept, on either side, the square grows larger eventually becoming much larger than the straight line.

The result must be continuous, smooth and monotonic on each side of the vertex and so appears parabolic.

The above principles applied to the square of the third degree polynomial $y = x(x - 1)(x - 2)$ lead to the graph of $y = x^2(x - 1)^2(x - 2)^2$ displayed in Figure 16.

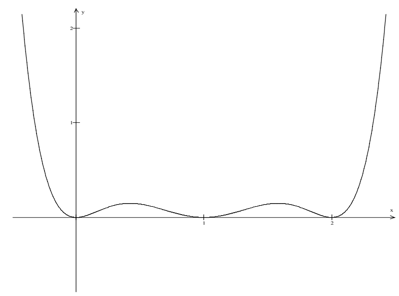


Figure 16 Squaring a Cubic

An interesting example is the square of the sine function. Because of the half wave symmetry of sine (x), that is, $\{\sin(x + \pi)\} = -\{\sin(x)\}$, it follows that $\{\sin(x + \pi)\}^2 = \{\sin(x)\}^2$. The graph of the squared function is a continuous smooth curve, which oscillates vertically at twice the frequency between values y = 0 and y = 1. The zeros of the squared function are all minima, tangential to the horizontal axis. The maxima of the square occur at the horizontal values of the extrema of the original sinusoid. The graphs of $y = \sin(x)$ and $w = \{\sin(x)\}^2$ are shown in figures17 and figure 18.

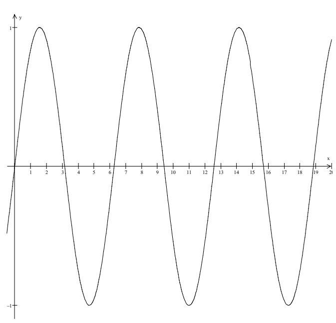


Figure 17 $y = \sin(x)$

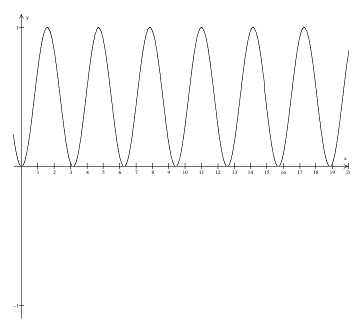


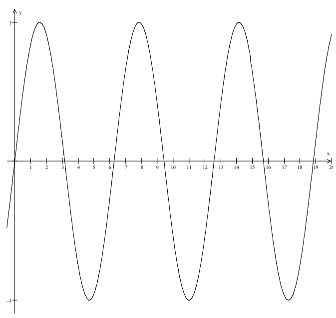
Figure 18 $w = \{\sin(x)\}^2$

Note should be made that the squared sine can be constructed by subtracting a cosine with half the amplitude and twice the frequency from the constant $y = \frac{1}{2}$, that is,

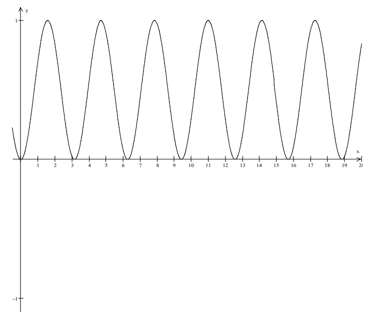
$\{\sin(x)\}^2 = \frac{1}{2} - \frac{1}{2}\cos(2x)$. This construction provides graphical confirmation of the double angle formula:

$$\cos(2x) = 1 - 2\{\sin(x)\}^2.$$

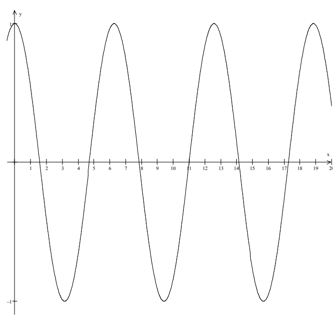
Similarly the graphs $y = \{\sin(x)\}^2$ and $y = \{\cos(x)\}^2$ dovetail as shown in figure 19. At every x the sum of these functions equals the constant, one, providing a visual confirmation of the Pythagorean Theorem. This graphing technique can provide a visual confirmation for any mathematical identity of smooth functions.



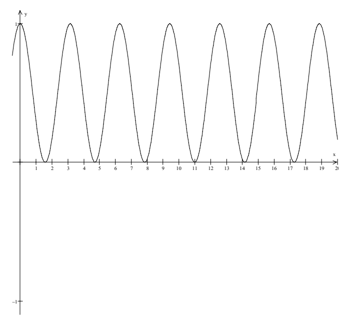
$y = \sin(x)$



$w = \{\sin(x)\}^2$



$y = \cos(x)$



$w = \{\cos(x)\}^2$

Figure 19 Graphical confirmation of the Pythagorean Theorem

Cubing

The principles relating to cubing a function are:

- The cube of any function preserves the signs of the function on each interval.
- The cube of a differentiable or continuous function will be differentiable or continuous.
- The cube of a monotonic function is monotonic.
- If a function has a maximum (minimum), the cube of the function will have a maximum (minimum) at the same horizontal value.
- The cube of the linear function $y = m(x - h)$ crosses the x-axis with a horizontal tangent at $x = h$. The cube of any function with a linear zero has a horizontal tangent at the x-intercept.

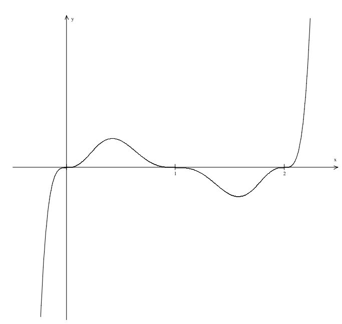


Figure 20
Cubing a polynomial

As an example the above principles can be applied to the cube of the third degree polynomial $y = x(x - 1)(x - 2)$ and will lead to the graph, shown in figure 20 of the function, $y = x^3(x - 1)^3(x - 2)^3$.

Taking the reciprocal

The principles relating to taking the reciprocal function are:

- Reciprocals of small but not zero numbers are large and vice versa.
- One and negative one are the only numbers which equal their reciprocals.
- Reciprocals of positive (negative) numbers are positive (negative).
- At the location of a zero of a function, the reciprocal will have a vertical asymptote.
- The reciprocal of the reciprocal of a function is the original function.
- The reciprocal of a function with a zero of odd degree has a pole that changes signs.
- The reciprocal of a function with a zero of even degree has a pole, which does not change sign.

The principles can be observed to apply to the reciprocals of tilted straight lines, which yield hyperbolas, asymptotic to the horizontal axis and with poles at the zeros of the straight lines. The reciprocals of quadratic functions representing vertical parabolas produce different curves depending on whether or not the parabolas cross the x-axis. Imagine that the vertex of an upward going parabola with no roots is lowered until the vertex touches the x-axis and then is lowered further producing two roots as shown in figure 21.

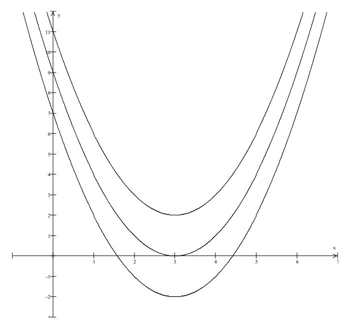


Figure 21 Lowering a parabola

The resulting reciprocals are graphed in figures 22, 23 and 24. As the vertex is lowered the peak of the reciprocal rises and becomes infinite when the vertex touches the x-axis. As the roots appear the reciprocal exhibits vertical asymptotes that separate as the roots move apart.

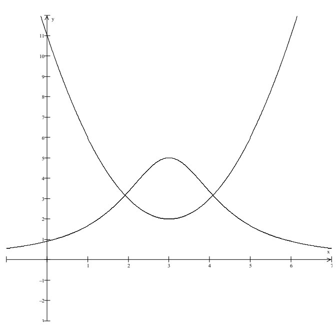


Figure 22
$y = (x-3)^2 + 2$
and its reciprocal
(not to scale)

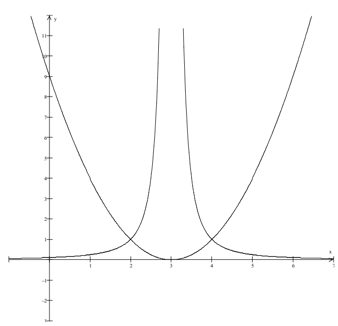


Figure 23
$y = (x-3)^2$
and its reciprocal

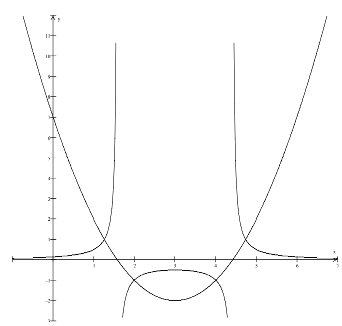


Figure 24
$y = (x-3)^2 - 2$
and its reciprocal

The graphs of the cubic polynomial, $y = x(x-1)(x-2)$ and its reciprocal
$y = \dfrac{1}{x(x-1)(x-2)}$ are shown in figure 25.

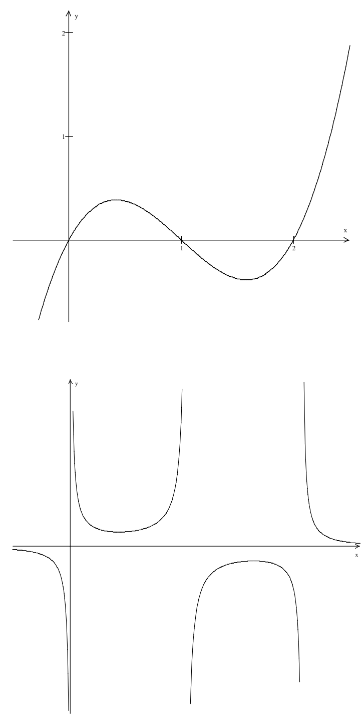


Figure 25 A cubic and its reciprocal

Let us examine general rational functions, quotients of two polynomials. Those rational functions, called proper, whose numerator polynomial degree is less than their denominator degree, are asymptotic to the horizontal axis. Improper rational functions can be divided to yield the form of the sum of a polynomial and a proper rational function. This polynomial will determine the remote behavior of the original improper rational function. In the cases when the numerator and denominator polynomials can be completely factored, the zeros and poles and their degrees can be found which will determine the near behavior. Thus the determination of these features facilitates sketching a more general set of rational functions.

Multi-valued algebraic curves, fractional powers and approaches to the x-axis.

To begin a study of algebraic curves, one might want to examine the simple expressions of the form $y = x^p$; where p is a rational number. We have already examined the cases where p is a positive integer including: the identity function, $y = x$, the square, $y = x^2$, and the cube $y = x^3$. If p is a negative integer the function can be treated as the reciprocal of a case where p is positive. Fractional powers will be studied next. The identity function, where $p = 1$, is simply a diagonal line crossing the x-axis at an angle. If $p > 1$ the curves all approach the x-axis with a zero slope. When single-valued even roots are taken, the curve appears only in the first quadrant.

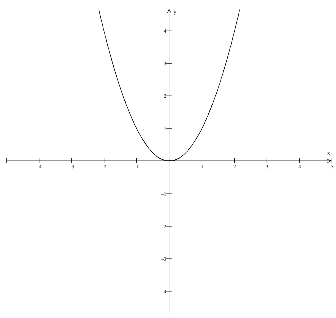


Figure 26

$p = 2$

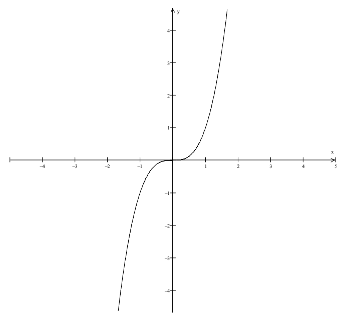


Figure 27

$p = 3$

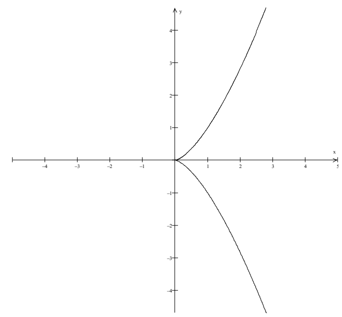


Figure 28

$p = \frac{3}{2}$

But it is seen that for different values of p, the approaches to the x-axis can be different. It is possible for the curve to:

1) not to cross the x-axis but approach it with a zero slope, touch it and return in the direction from which it came such as the parabola in figure 26 where $p = 2$,
2) be single-valued and cross the x-axis with a zero slope such as the cubic where $p = 3$ as in figure 27 or
3) approach the x-axis with a zero slope and double back on itself like the semi-cubic parabola in figure 28 where $p = \frac{3}{2}$.

In the cases where $0 < p < 1$ all the curves approach the x-axis with in infinite slope. And again, it is seen that for different fractions, the approaches can be different.

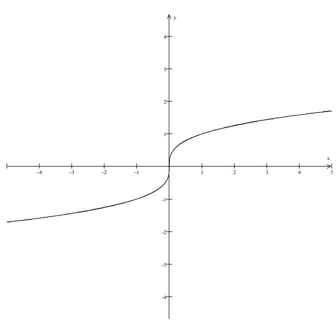


Figure 29

$p = \frac{1}{3}$

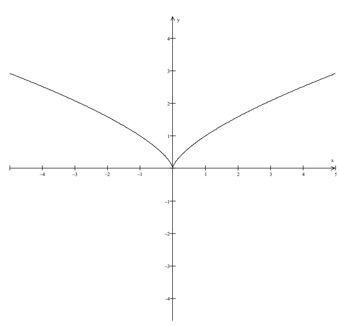


Figure 30

$p = \frac{2}{3}$

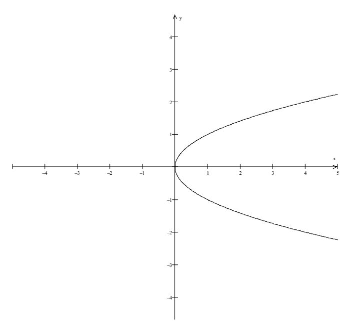


Figure 31

$p = \frac{1}{2}$

It is possible for the curve to:

1) be single-valued and cross the x-axis with a vertical slope such as the curve of the cube root in figure 29 where $p = \frac{1}{3}$,

2) not cross the x-axis but to have a cusp with an infinite-slope at the point of contact such as the curve in figure 30 where $p = \frac{2}{3}$ or

3) double back on itself with a vertical slope as does the horizontal parabola in figure 31 where $p = \frac{1}{2}$.

If $0 < p = \frac{n}{m} < 1$, the principles governing these behaviors can be stated.

- If m and n are both odd, the curve crosses the x-axis with a vertical slope.
- If m is odd and n is even, the curve touches the x-axis as a vertical cusp.
- If m is even and n is odd, the double-valued root crosses the x-axis with a vertical slope doubling back on itself like the horizontal parabola; the single-valued root appears only in the first quadrant.
- If n and m are both even, the curve can have a double vertical cusp symmetric about the x-axis, like two horizontal parabolas with a common vertex and whose axis of symmetry is the x-axis. More will be said later.

If p is negative then the curve has a pole.

Fractional powers, taking square roots

The principles relating to taking the square root of a function are:

- Intervals on which the arguments of the square roots are negative are excluded.
- Positive numbers may have both positive and negative square roots.
- If $0 < y < x$, then $|\sqrt{y}| < |\sqrt{x}|$.
- The square root of the straight line, the function $y = m(x - a)$, will be a horizontal parabola, defined on the infinite interval where y is positive, whose axis of symmetry is the x-axis and whose vertex is at the zero, $x = a$.
- Employing the previous principle we note that the square root of a function that has a linear zero at $x = a$ will cross the x-axis at $x = a$, doubling back on itself with an infinite slope.
- If a positive function has a maximum at $x = a$, then the positive branch of the square root of the function also will have a maximum at $x = a$, while the negative branch will have a minimum at $x = a$.

The square of a square root and the square root of a square

Restricted to the domain of positive numbers, for a single-valued square root, both of the functions $y = (\sqrt{x})^2 = (x^{1/2})^2 = x$ and $y = \sqrt{(x^2)} = (x^2)^{1/2} = \pm x$ have the same graph. But there are viewpoints from which the graphs of the equations apparently differ.

1) If the straight line $y = x$ is squared first and then the single-valued square root is taken, the resulting graph is the absolute value function, $y = |x|$ as is shown in figure 33.
2) If the square root is taken first the result is excluded for negative x. Squaring then produces only the ray in the first quadrant as shown in figure 32.
3) If the straight line $y = x$ is squared first and then the double-valued square root is taken, which is appropriate for graphing the equation $y^2 = x^2$, the resulting graph as shown in figure 34 is the union of the two diagonal straight lines $y = \pm x$.

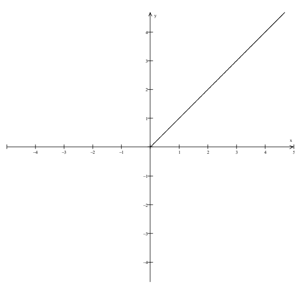


Figure 32 $y = (\sqrt{x})^2$

either single or double-valued square root

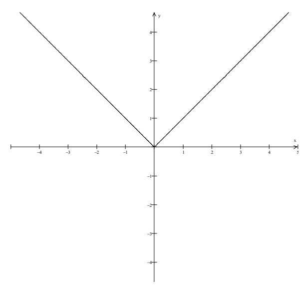


Figure 33 $y = \sqrt{x^2}$

single-valued square root

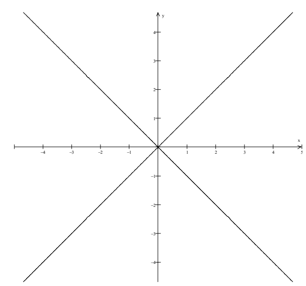


Figure 34

$y = \pm\sqrt{x^2} = \pm(x^2)^{1/2}$

double-valued square root

This means the function operations, square root and squaring, are not commutative. Squaring is always single-valued and produces no negative values. The double-valued square root operation is not defined for negative values of x but produces both plus and minus values for the positive values of x.

Principle: The double-valued square root of a double zero crosses the x-axis like an × as in figures 34 and 39. The slopes of the double approach to the x-axis will be finite, non-zero and equal in absolute value.

Returning to the before-mentioned concern of $y = x^p$, $p = \frac{n}{m} > 0$, where both n and m are even, allowing for double-valued square roots, consider the cases: $p > 1$, $p = 1$ and $0 < p < 1$. If the root is performed first the curve will only appear in the first quadrant. If the power is performed first then the curve will appear in all four quadrants as is shown in figures 34, 35 and 36.

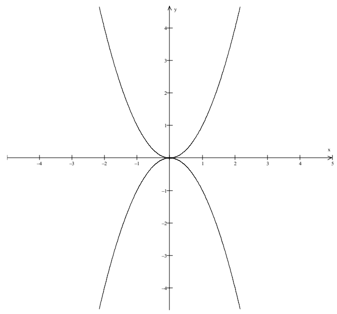


Figure 35

$0 < p = \frac{n}{m} < 1,$

both n and m even,
performing power before
taking the double root.

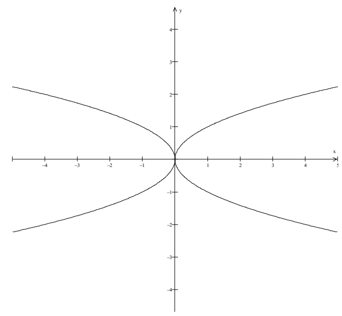


Figure 36

$1 < p = \frac{n}{m}$,

both n and m even,
performing power before
taking the double root.

The difference is that when p > 1 the tangent line is horizontal; when 0 < p < 1, the tangent line is vertical. There is merit for engineers to sketch the entire multi-valued curve. If the problem at hand only needs a portion of the curve it is easy to discard the irrelevant part of the curve. Automatically discarding part of a curve may lose viable solutions.

The mathematics community has invested a great deal of intellectual effort in defining the basic concept of calculus, the function, as an inherently single-valued entity. Perhaps consideration should be given to making calculus the study of finitely multi-valued, algebraic curves. In such a study a student could contemplate tilted parabolas and circles in their entirety. The algebraic curves have much in common with single-valued functions including slopes and little in common with other general multi-valued relations such as those described by inequalities, like $y < x$.

Consider the effect of taking the square root of various quadratic functions as shown in figures 37 and 38. Lower an upward opening parabola. Initially if the vertex is above the x-axis, the square root is a double-valued hyperbola with values for every x. Continue lowering the parabola and the vertices approach each other, maintaining the same asymptotes. When the parabola touches the x-axis, the graph becomes the union of two straight lines, which are the asymptotes. Continue moving the parabola downwards and the graph again becomes a double-valued hyperbola but now with an excluded interval. Throughout the lowering process, as is shown in figure 37, the asymptotes remain unchanged.

Figure 38 displays the effect of taking the square root of a downward opening parabolic function, which is being lowered. One observes an ellipse, which diminishes until it becomes a point and then disappears entirely. Table 1 records the results of taking the square root of all of the combinations of quadratic functions.

	Parabola	Conditions	Result of taking the square root
1	$y = -a(x - h)^2 + k$	$a > 0,\ k > 0$	an ellipse whose domain lies between the zeros of the parabola
2	$y = -a(x - h)^2$	$a > 0$	only the single point (h, 0)
3	$y = -a(x - h)^2 - k$	$a > 0,\ k > 0$	no points satisfy the conditions
4	$y = a(x - h)^2 - k$	$a > 0,\ k > 0$	A double-valued hyperbola with an excluded interval between the zeros of the parabola
5	$y = a(x - h)^2 + k$	$a > 0,\ k > 0$	A double-valued hyperbola defined everywhere
6	$y = a(x - h)^2$	$a > 0$	two intersecting diagonal lines; the asymptotes of the above hyperbolas

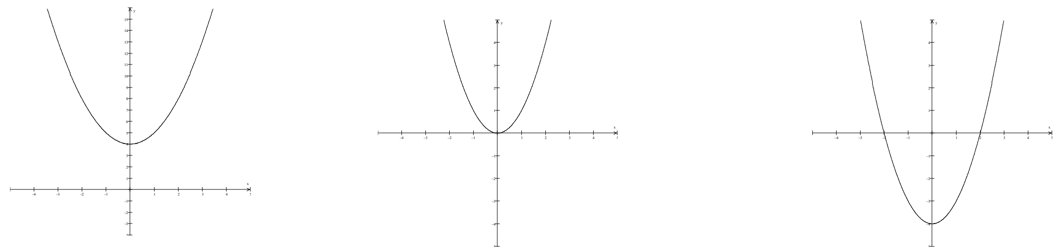
Table 1 Taking the Square Root of a Quadratic Function

A pretty example, shown in Figure 39, illustrates the power of visual analysis. The graph of the quartic polynomial $y = (x + 4)(x - 1)^2(4 - x)$ is displayed in the upper half of figure 39. The double-valued graph of its square root, $y = \pm (x - 1)(16 - x^2)^{1/2}$, can be easily obtained and is shown in the lower half of figure 39. The self intersection at $x = 1$ is a signature of a square root of a square. The infinite intervals where $|x| > 4$ are seen to be excluded.

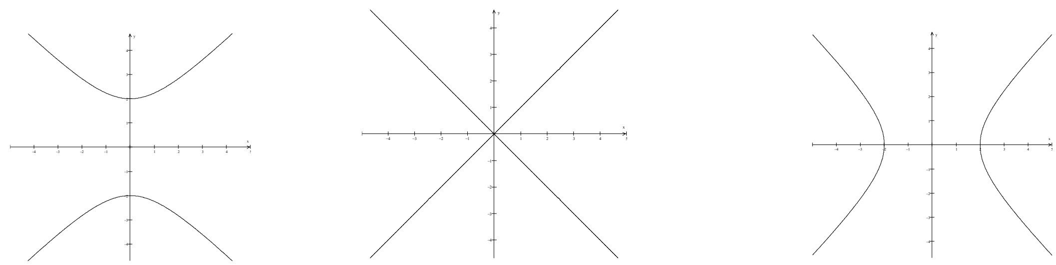
Fractional powers, taking cube roots

The principles relating to taking the cube root of a function are:

- All numbers have cube roots.
- The cube roots of positive (negative) numbers are positive (negative).
- Since the cube root function is monotonic, then if $x < y$, then $\sqrt[3]{x} < \sqrt[3]{y}$.
- The cube root of function that has a linear zero at x = a will cross the x-axis perpendicularly at x = a.
- The cube root of function that has a second degree zero at x = a will not cross the x-axis but will have a cusp with an infinite slope at x = a.

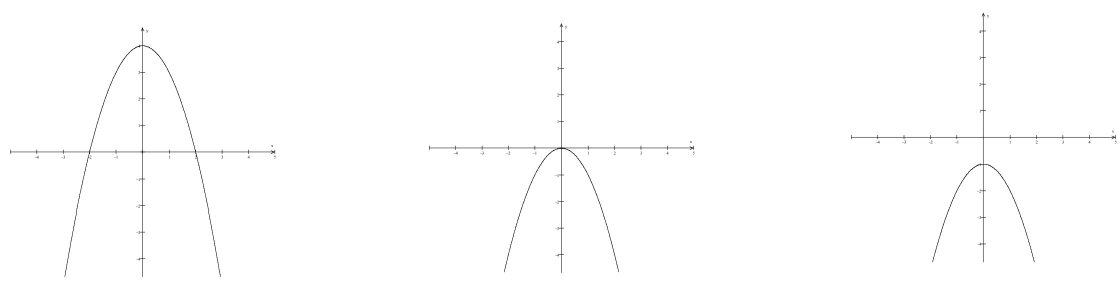


$y = x^2 + 4$ $\qquad$ $y = x^2$ $\qquad$ $y = x^2 - 4$

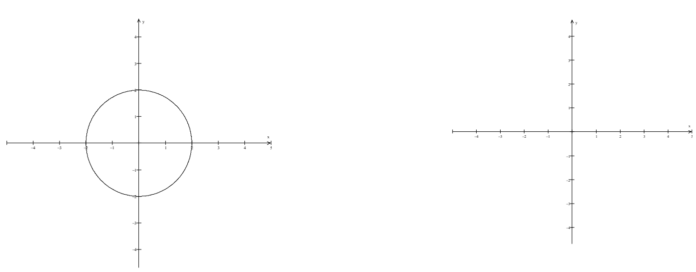


$y = \sqrt{x^2 + 4}$ $\qquad$ $y = \sqrt{x^2}$ $\qquad$ $y = \sqrt{x^2 - 4}$

Figure 37 Upward opening quadratic functions and their square roots



$y = 4 - x^2$ $\qquad$ $y = -x^2$ $\qquad$ $y = -4 - x^2$



$y = \sqrt{4 - x^2}$ Usually an ellipse

$y = \sqrt{-x^2}$ Only one point

$y = \sqrt{-x^2 - 4}$ Excluded graph

Figure 38 Downward opening quadratic functions and their square roots

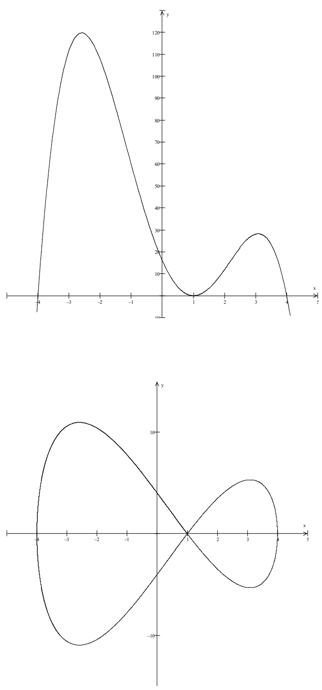


Figure 39
The square root of a quartic

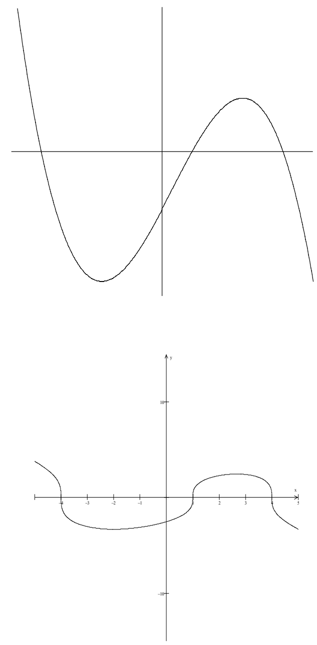


Figure 40
The cube root of a cubic

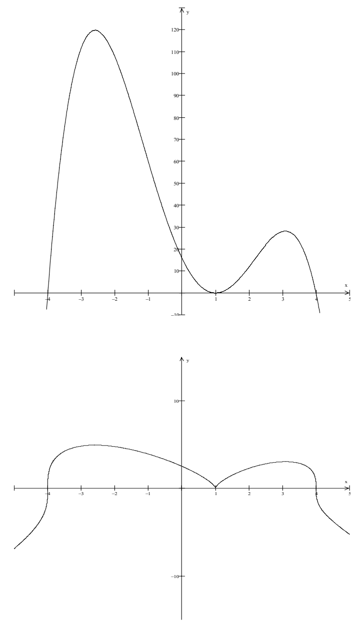


Figure 41
The cube root of a quartic

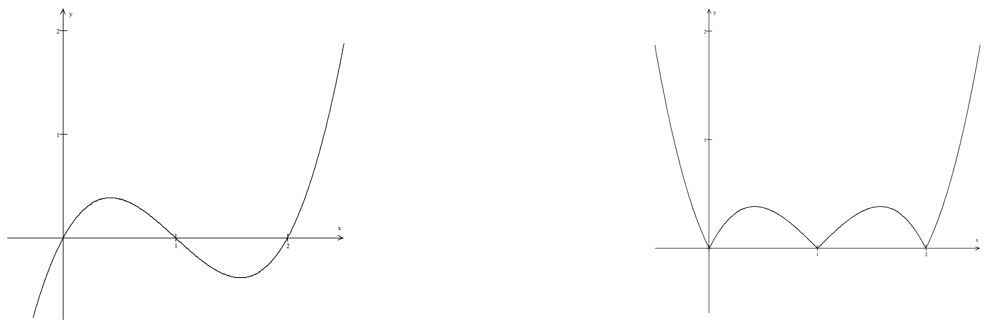
The cube root of the cubic with linear roots $y = (x + 4)(x - 1)(4 - x)$ is shown in figure 40. The curve is seen to cross the x-axis perpendicularly.

The quartic polynomial $y = (x + 4)(x - 1)^2(4 - x)$ has a double root at $x = 1$ but only single roots at $x = \pm 4$ and so its cube root must have a vertical cusp at $x = 1$ as can be seen in figure 41. The curve is seen to cross the horizontal axis perpendicularly at $x = \pm 4$.

Taking the absolute value

During the early 1900's electrical engineers needed circuits that would convert sinusoidal electrical energy to DC constant voltage. These circuits were called rectifiers. One type of rectifier, called a half wave rectifier, would simply clip the negative part of the sine wave but retain the positive part of the signal, which would then be filtered to reduce the "ripple." Another type of rectifier, called a full wave rectifier, would flip the negative part of the sine wave making it positive. The output of these full wave rectifiers was what mathematicians call the absolute value of the sine wave, having a DC component and a "ripple" of twice the original frequency. While integrals of functions that change sign on an interval may not be directly related to any area, the integral of the absolute value of a function is sum of the area under the positive part of the signal and above the x-axis added to the area below the x-axis but above the negative part of the signal.

The operation of absolute value can be viewed as flipping the curve on the intervals where the function is negative and leaving the curve unmodified where the function is positive. This is equivalent to taking the single-valued square root of the square of the function. The graphs of the cubic polynomial $y = x(x - 1)(x - 2)$ and its absolute value are compared in figure 42.



$y= x(x - 1)(x - 2)$ $\qquad$ $y = |x(x - 1)(x - 2)|$

Figure 42 Comparison of a cubic and its absolute value

7. Graphical oddities

After a course in calculus many students are left with the impression that the graphs of equations are always curves. To correct the record, the following examples display equations whose graphs are sets of discrete points, unions of common curves and even closed regions containing area in the xy−plane.

Equations for the unions and intersections of curve

Say the equations for two curves in implicit form are: $F(x, y) = 0$ and $G(x, y) = 0$. Every point in the intersection of the two curves satisfies the equation;

$$\{F(x, y)\}^2 + \{G(x, y)\}^2 = 0.$$

Every point in the union of the curves satisfies the equation; $F(x, y)\, G(x, y) = 0$.

Equations for the regions inside and including the unit circle

Usually the solution of a conditional equation in one variable is satisfied by only a discrete set of values. However, the solution of the following conditional equation in x consists of the entire interval $0 \le x \le 1$.

$$|x| + |1 - x| - 1 = 0$$

This construction works because the sum of the two absolute value functions produces a piecewise-defined function, which is constant on the interval between the vertices at $x = 0$ and $x = 1$. The function, $y = |x| + |1 - x| - 1$ is zero only on the interval $0 \le x \le 1$.

This feature can be exploited in finding a polar equation that represents a region in the two dimensional plane. If r represents polar distance from the origin of a plane, an equation satisfied by any point in or on the unit circle is:

$$|r| + |1 - r| - 1 = 0$$

Equations for the region inside of and including the unit square.

Applying the previous principle we find every point inside and on the boundary of the unit square $0 \le x \le 1$, $0 \le y \le 1$ satisfies the equation:

$$(|x| + |1 - x| - 1)^2 + (|y| + |1 - y| - 1)^2 = 0$$

Although the above equation was desired, the following equation, which has an interesting graph, was mistakenly typed. This graph is left as an exercise.

$$(|x| + |1 - x| - 1)(|y| + |1 - y| - 1) = 0$$

8. Classic examples of visualizations and Euler's constant, γ

It is of value to recognize a few classical visual demonstrations that were known by the mathematicians of the past. While mathematicians today might not accept these pictures as proofs, the pictures must have guided the thinking of past mathematicians. Every functional identity states that two apparently different functions are in reality different forms of the same function and therefore possess the same table and the same graph.

Courant's calculus text [3] contains the following visual interpretation of the equation for integration-by-parts. Imagine a rectangle in the first quadrant whose sides are parallel to the axes. Draw the continuous curve of a monotonically increasing function connecting the lower left corner A with the upper right corner B. The integration-by-parts formula states that the area of the rectangle determined by B and the origin equals the sum of areas under the curve between A and B, the area to the left of the curve between A and B and to the right of the y-axis and the area of the rectangle determined by the origin and A. This picture should be shown to all engineering calculus students. Also on the subject of integration, the well-known visual demonstration should be noted of the convergence on a finite interval of the Riemann integral for piecewise monotonic functions.

Euler's Constant

Euler's constant, γ, can defined visually, as shown in figure 43, as the limit of the sum of the areas of the infinity of curved triangles that lie below the descending staircase, $y = \frac{1}{floor(x)}$ but above the monotonically decreasing hyperbola $y = \frac{1}{x}$: $x \ge 1$. The area below the staircase function and the area below the hyperbola are both infinite, but the limit of the sum of the areas of the curved triangles between them can be shown visually to be less than one. All that needs to be done is to slide all the triangles (their base equals one) to the left so that they lie over the interval $0 \le x \le 1$. They are all seen to fit inside a square of side of one, so the infinite sum, γ, must be less than one. If the vertices of the staircase were connected with straight lines the limit of the sum of the areas of the resulting straight triangles will be 0.5 . (This can be seen in the picture or derived by telescoping the series) But since the hyperbola is concave up, the limit of the sum of the areas of the curved triangles must be a somewhat larger than 0.5, which is what Euler ingeniously calculated, approximately 0.577. What is so clear in a picture takes effort both to describe and to read in words and algebra.

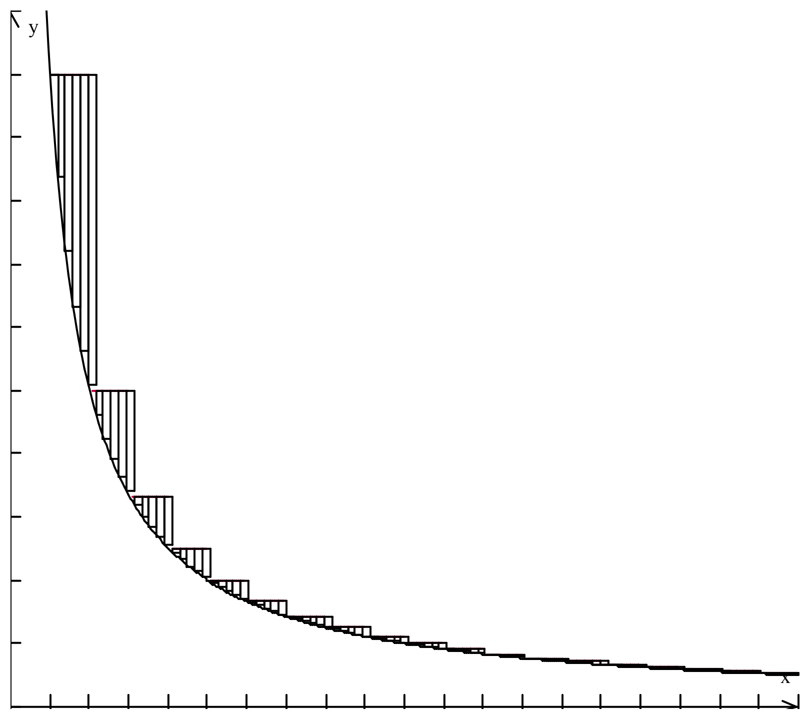


Figure 43 Euler's Constant: The sum of the areas of the triangles between the functions $y = \dfrac{1}{floor(x)}$ and $y = \dfrac{1}{x}$. (not to scale)

9. Conclusion

This paper places the spotlight on three mathematical areas:

1) Visual analysis and curve sketching
2) The arithmetic and functional principles that provide the basis of the graphical constructions.
3) Composition of functions as a functional mechanism.

Visual techniques are marvelous, insightful, practical and just plain fun. Engineers commonly use them in seeking insight and solutions to problems of analysis and design. There can be a benefit to a student who can acquire the ability to visualize equations. Today, the mathematics community tolerates visual insights, but it must be admitted that the mathematicians of the mid-twentieth century erred when they disparaged visual approaches. It was wrong to deny the students of that era the common sense insight that comes with visualization.

The principles are difficult to write and difficult to read but once understood, easy to either remember or reconstruct. A high school student could already be acquainted with many of them, but perhaps in a context unrelated to curve sketching. This paper provides an opportunity for an instructor to either review or introduce the principles whichever is appropriate for an individual student. The placement of the principles adjacent to the graphs may assist a student in seeing an immediate application in curve sketching. A student does not have to learn all the principles at one time but should be aware of them concurrently with or preceding the study of calculus.

Composition of functions is an important concept in the construction of functions. Composition should not be treated as an appendage to either "related rates" or the chain rule. As a concept it stands on its own and must precede considerations of continuity and differentiability.

Many marvelous examples of visualizations and principles have been explored. They clearly indicate what could be achieved if techniques of visualization are introduced early into the mathematics curriculum and in classrooms.

Bibliography:

1. Aleksandrov, et al "Mathematics Its Content, Methods and Meaning," The MIT Press, Cambridge MA, 1963
2. Alsina, Claudi & Nelson, Roger B. "Math Made Visual," MAA, Washington, 2006
3. Courant, R. "Differential and Integral Calculus" Volume I, Interscience Publishers Inc. NYC, NY. 1937
4. Dunham, William, "Euler, the Master of Us All," MAA, Washington, 1999 p 35
5. Grossfield, Andrew "Wonder, Discovery and Intuition in Elementary Mathematics," Proceedings of the ASEE Zone 1 Conference at West Point, 2008
6. Nelson, R. B. "Proofs without Words: Exercises in Visual Thinking," MAA Washington, 1993
7. Nelson, R. B. "Proofs without Words II: More Exercises in Visual Thinking," MAA Washington, 2000