

# Advanced Math: Notes on Lessons 118-121

David A. Wheeler, 2008-04-07

## Lesson 118: Roots of Polynomial Equations

(This lesson is essentially a continuation of 117B, the rational root theorem)

Sometimes you can solve higher-degree polynomial equations through two steps: (1) use the rational root theorem (lesson 117) to find *some* of the roots, (2) divide the original equation by each of the (*x-root*) values that you found, and then (3) solve what's left over.

For example, given:

$$x^3 - 10x^2 + 33x - 34 = 0$$

The rational root theorem says that if there are rational roots, they must be in this set:

$$\frac{\pm 1, 2, 17, 34}{1}$$

If  $x=1$ ,  $x^3 - 10x^2 + 33x - 34 = 1^3 - 10(1)^2 + 33(1) - 34 = -10$ , which isn't zero, so it's not a root.

If  $x=2$ ,  $x^3 - 10x^2 + 33x - 34 = 2^3 - 10(2)^2 + 33(2) - 34 = 8 - 40 + 66 - 34 = 0$ , so it's a root!

This means that you can factor the original polynomial with  $(x-2)$ . Since you can, divide the original polynomial by  $(x-2)$ , and you end up with:  $x^2 - 8x + 17$ . Now it's down to second-degree, which you know how to solve.. just use the quadratic equation:

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

and you'll find  $x = 4+i, 4-i$ .

So the entire factored polynomial is:

$$x^3 - 10x^2 + 33x - 34 = (x-2)(x-4-i)(x-4+i)$$

and the original equation is the same as:

$$(x-2)(x-4-i)(x-4+i) = 0$$

## Lesson 119: Descartes' Rule of Signs / Upper and Lower Bound Theorem / Irrational Roots

### Descartes' Rule of Signs

If you write a polynomial in its usual unfactored form (sorted from the largest to smallest exponent), you can quickly count the number of "sign changes" in the nonzero coefficients. For example, in:

$$3x^6 - 3x^5 + 2x^4 - 5x^2 - x + 4 = 0$$

there are 4 sign changes: from  $3x^6$  to  $-3x^5$ , from  $-3x^5$  to  $+2x^4$ , from  $+2x^4$  to  $-5x^2$ , and from  $-x$  to  $+4$ . Notice that the zero-valued coefficients (such as the missing  $x^3$ ) are irrelevant for this rule.

Descartes' rule of signs says that "the number of positive real roots of a polynomial equation with real

coefficients equals the number of sign changes minus 0 or a positive even number (2, 4, ...)."

You can also use this to find the largest possible number of negative real roots by calculating  $p(-x)$ , i.e., replacing all the "x"s with "-x". This has the effect of flipping the sign of all the coefficients with odd exponents.

For example, if  $p(x) = 3x^6 - 3x^5 + 2x^4 - 5x^2 - x + 4$ , then  $p(-x) = x^6 + 3x^5 + 2x^4 - 5x^2 + x + 4$ .

Thus, Descartes' rule of signs also says that "the number of negative real roots of a polynomial equation  $p(x)$  with real coefficients equals the number of sign changes in  $p(-x)$  minus 0 or 2 or 4 or ...".

### **Upper and Lower bound theorem**

The "region of interest" approach (lesson 116) sometimes doesn't give enough information on where the zeros of the polynomial (the roots of the equation) are; the "upper and lower bound theorem" may be able to give a narrower answer. This only works if the polynomial has real coefficients.

For an "upper bound", repeatedly pick positive values (usually you'd use integers) and use synthetic division (lesson 113). If the nonnegative numbers in the bottom row are *all*-positive or *all*-negative, then there is no real root larger than that divisor. **Beware:** if both positive and negative values result, it's *still* possible that there is no real root larger than that divisor. In other words, the inverse isn't necessarily true. (The Saxon book doesn't make that clear at all.)

You'll need to do a little searching to find the integer  $x$  where this is true, and where it's not true for  $x-1$ ; at that point,  $x$  is the smallest integer upper bound that this theorem can find.

Let's try this with the following polynomial:

$$x^4 - 3.5x^3 - 37x^2 + 42.5x + 75$$

First, let's see if 4 is larger than all the roots, using synthetic division:

<u>4</u>	1	-3.5	-37	42.5	75
		4	2	-140	-390
	1	0.5	-35	-97.5	-315

This didn't work; we have both positive and negative numbers. Note the the final bottom-right value is nonzero, so 4 isn't a root either. Let's try again, with 8:

<u>8</u>	1	-3.5	-37	42.5	75
		8	36	-8	276
	1	4.5	-1	34.5	351

We still have both positive or negatives, but they're fewer - we may be getting close. Now, it turns out that in this example, the largest zero of this polynomial is 7.5, and so 8 actually *is* larger than any zero. This theorem *didn't* catch that. I want to emphasize this because the Saxon book doesn't make it clear: There's no guarantee that this theorem will find the *smallest* real-life integer upper bound, just that it will find *some* bounds.

Continuing with our example, 8 didn't work, so let's try 9:

<u>9</u>	1	-3.5	-37	42.5	75
		9	49.5	112.5	1395
	1	5.5	12.5	155	1470

We know 9 is not a zero, because we have a nonzero remainder in the synthetic division. But we have all-positive values at the bottom, so we know for sure that all the (real) zeros have values of 9 or less. We didn't get that result with 8, so among the integers, 9 is the smallest upper bound we can find using this method. And since 9 isn't a zero, we know that all real zeros are less than 9.

There's a similar rule for a lower bound. Again, use synthetic division, but use negative numbers (usually integers) as the synthetic divisors. A divisor is smaller than any real root if the synthetic divisor's bottom row has alternating signs; zero can be considered "either sign" for this purpose.

Let's try -4:

<u>-4</u>	1	-3.5	-37	42.5	75
		-4	30	28	-282
	1	-7.5	-7	70.5	-207

This doesn't alternate; -7.5 is followed by -7, and both are negative. Let's try -6:

<u>-6</u>	1	-3.5	-37	42.5	75
		-6	57	-120	465
	1	-9.5	20	-77.5	540

This does alternate - which means that there are no roots smaller than -6. But perhaps we can find a smaller value that will work - we haven't tried -5 yet. So let's try that:

<u>-5</u>	1	-3.5	-37	42.5	75
		-5	42.5	-27.5	-75
	1	-8.5	5.5	15	0

We have a remainder of 0, so this is a zero!! We don't have alternating signs, so we don't know for sure if there are smaller zeros or not. So this approach would report that the zeros are between -6 (lower bound) and 9 (upper bound).

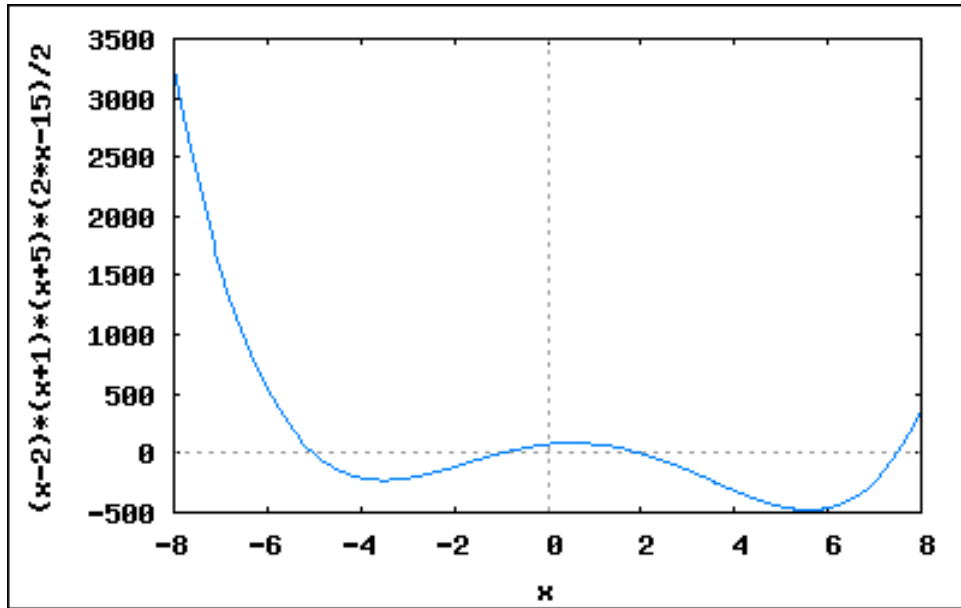
By the way, I created this example by first deciding where I wanted the roots; once I'd done that, I then wrote the polynomial in factored form:

$$(x-7.5)*(x-2)*(x+1)*(x+5)$$

I then multiplied it out, which is of course:

$$x^4 - 3.5x^3 - 37x^2 + 42.5x + 75$$

Here's a graph of this particular polynomial:



You *can't* use the “rational roots theorem” (lesson 117) directly to find all the rational roots of:

$$x^4 - 3.5x^3 - 37x^2 + 42.5x + 75 = 0$$

because the “rational roots theorem” requires that all of the polynomial coefficients be *integers*. Clearly, -3.5 and 42.5 are not integers! You can use the rational roots theorem, though; simply multiply both sides by some value that turns them all into integers. In this case, multiplying both sides by 2 will do it:

$$2x^4 - 7x^3 - 74x^2 + 83x + 150 = 0$$

All of the rational roots will have this form:

$$\pm \frac{\text{factors of } 150}{\text{factors of } 2}$$

Since  $150 = 2 \times 3 \times 5 \times 5$ , we have a lot of possible factors, and these *do* cover all the roots:

$$\pm \frac{\{1, 2, 3, 5, 6, 10, 15, 25, 30, 50, 75, 150\}}{\{1, 2\}}$$

As discussed in lesson 118, once you've found two of the roots, you can divide the 4th degree polynomial by  $(x - \text{root})$  with each of those roots. The result will be a second-order polynomial, which you can solve directly using the quadratic equation. Eventually you'd find the roots  $x = -5$ ,  $x = -1$ ,  $x = 2$ , and  $x = 7.5$ .

### ***Irrational roots***

If a real root is irrational (can't be expressed as a fraction of two integers), there's no simple way of finding its value in general. However, if a polynomial evaluates to a positive value at some value  $x = a$ , and to a negative value at  $x = b$ , then clearly it evaluated to zero somewhere between  $a$  and  $b$ . That's

because polynomials are always continuous, and zero is between the positive and negative values. The same is true the other way; if  $x=a$  is negative, and  $x=b$  is positive, then again there's a zero of that polynomial between  $a$  and  $b$ .

## Lesson 120: Matrix Algebra / Finding Inverse Matrices

If  $A$  and  $B$  are real numbers, then given the equation:

$$Ax = B$$

You can solve this by multiplying both sides by  $A^{-1}$ , which is just  $1/A$ :

$$A^{-1}Ax = A^{-1}B$$

Since  $A^{-1}A = 1$ , that means that:

$$x = A^{-1}B$$

The same is true when  $A$  and  $B$  are matrices, and “ $x$ ” is replaced with a matrix that is a column of variables. In fact, this is one of the primary uses of matrices! By doing this, you can solve a massive number of simultaneous linear equations with a huge number of variables. When they get large, you'll want to do this by computer, not by hand, but we can learn how to do the 2-variable case, which only involves 2x2 matrices at most.

Let's first examine a simple pair of linear equations:

$$2x + 3y = 4$$

$$5x + 6y = 9$$

This can be quickly rewritten as matrix multiplication:

$$\begin{bmatrix} 2 & 3 \\ 5 & 6 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4 \\ 9 \end{bmatrix}$$

Now to solve this (we want to solve for the matrix of variables on the left-hand-side), we need to multiply both sides of the equation by the inverse of the 2x2 matrix. This special matrix is called the “inverse matrix”; any square matrix with a nonzero determinant has an inverse matrix.

Here's the inverse of any 2x2 matrix:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

The inverse is the  $1/\text{determinant}$  followed by a modified version of the original matrix. That's why the determinant has to be nonzero - if it were zero, we would divide by zero.

Except in special circumstances, matrix multiplication is *not* commutative: That is,  $A \cdot B$  is not always equal to  $B \cdot A$ . Thus, when “multiplying on both sides” with matrices, put them on the *same* position in both sides. Matrix multiplication *is* associative; that is,  $A \cdot (B \cdot C) = (A \cdot B) \cdot C$ , where  $A$ ,  $B$ , and  $C$  are matrices and/or complex numbers (including real numbers). Knowing that, let's solve the example:

$$\begin{aligned} \begin{bmatrix} 2 & 3 \\ 5 & 6 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} &= \begin{bmatrix} 4 \\ 9 \end{bmatrix} \\ \begin{bmatrix} 2 & 3 \\ 5 & 6 \end{bmatrix}^{-1} \begin{bmatrix} 2 & 3 \\ 5 & 6 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} &= \begin{bmatrix} 2 & 3 \\ 5 & 6 \end{bmatrix}^{-1} \begin{bmatrix} 4 \\ 9 \end{bmatrix} \\ \begin{bmatrix} x \\ y \end{bmatrix} &= \frac{1}{(2)(6)-(3)(5)} \begin{bmatrix} 6 & -3 \\ -5 & 2 \end{bmatrix} \begin{bmatrix} 4 \\ 9 \end{bmatrix} \\ \begin{bmatrix} x \\ y \end{bmatrix} &= \frac{-1}{3} \begin{bmatrix} 6 & -3 \\ -5 & 2 \end{bmatrix} \begin{bmatrix} 4 \\ 9 \end{bmatrix} \\ \begin{bmatrix} x \\ y \end{bmatrix} &= \frac{-1}{3} \begin{bmatrix} -3 \\ -2 \end{bmatrix} \\ \begin{bmatrix} x \\ y \end{bmatrix} &= \begin{bmatrix} 1 \\ 2/3 \end{bmatrix} \\ x=1, y=2/3 \end{aligned}$$

**Real-world warnings:** Matrix inversion is a very common operation, so it is built into a vast number of computer programs and calculators. Unfortunately, the “obvious” way of programming a matrix inversion sometimes gets terribly wrong answers. The problem is that most real numbers are only approximated by most software; these errors are usually so small that they don’t matter, but when calculating a matrix inversion the “obvious” way, these approximations can sometimes lead to terribly wrong answers. This is a solved problem – ways to compensate for this problem have been known for decades. Unfortunately, some programs don’t do use them. If you’re not sure you can trust the program that’s doing matrix inversions, check to make sure the compensating techniques have been used (software peer reviews and testing reduce this risk). In addition, even if the computer has compensating mechanisms, in some matrices a trivial variation in values will give a vastly different matrix inversion. No computer algorithm can compensate for that! In the “real world”, you’d also compute the matrix’s “condition number”; if it’s beyond about  $10^8$ , don’t use the matrix inverse at all, because the matrix inversion is too sensitive to the exact data values to be useful.

## Lesson 121: Piecewise Functions / Greatest Integer Function

### *Piecewise functions*

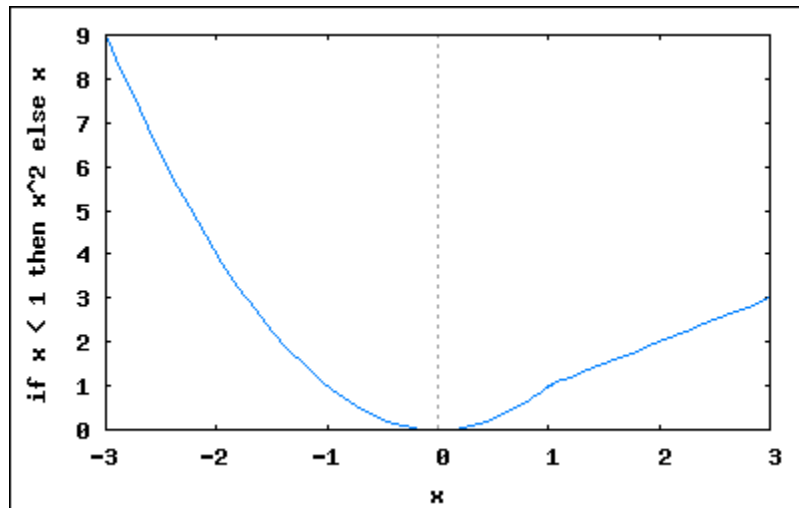
Real-world systems often don’t follow simple equations over all the domains we’re interested in, and there’s an easy way to handle that - we just define functions in pieces. Traditionally, this is shown with an opening curly brace on the left, followed by various expressions (one per line), and the expressions are then followed by conditions where that expression is valid. Some people omit the “if...” in a condition; it’s clearer if you include the word “if”, though, and note that Saxon does include the word “if”. For example:

$$f(x) = \begin{cases} x^2 & \text{if } x < 1 \\ x & \text{if } x \geq 1 \end{cases}$$

Many computer programs can handle piecewise functions, too, but many of them require that you give them in the opposite order, that is, you say “if” first, then the condition, and then the expressions depending on whether it is true or false. For example, in wxMaxima (a computer algebra system), the same function could be represented this way:

(if  $x < 1$  then  $x^2$  else  $x$ )

Here's a graph of this function:



Often these functions will be continuous, but in some cases there will be a “jump” from one position to another. Show these “jumps” on a graph using unfilled circles; the unfilled circle means “not here” (see the book for examples).

You should be able to create a graph from a piecewise function, and a piecewise function definition given the graph. Creating the functions from a graph should be easy: Just notice where there's a change from one function to another, and write those domains down (on the right-hand-side) first. Then, for each piece, figure out what the function is.

So if you started just from the graph above, it should be obvious that the change happens when  $x=1$ . While  $x \leq 1$ , we have a parabola, while when  $x \geq 1$ , we have a line. The graph seems connected (no “jump”), so it doesn't matter if you assign  $x=1$  to the parabola or the line, but you must assign it only once, and since  $x=1$  is graphed, you must assign it to *something*. Let's arbitrarily assign  $x=1$  to the line (we could assign it to the parabola instead). That means you'd have this:

$$f(x) = \begin{cases} \text{a parabola of some kind} & \text{if } x < 1 \\ \text{a line of some kind} & \text{if } x \geq 1 \end{cases}$$

Now we just need to fill in each of the equations; the result will be:  $f(x) = \begin{cases} x^2 & \text{if } x < 1 \\ x & \text{if } x \geq 1 \end{cases}$

## Greatest Integer Function

The greatest integer function is written using traditional mathematical notation as full square brackets around some expression, e.g.:  $[x]$ . In computer languages it's often written as `INT()` or `INTEGER()`, though there are many alternatives. If the expression is an integer, it returns the integer; otherwise, it returns the greatest integer less than the expression. Here are few examples that should clarify this:

$$[3] = 3$$

$$[2.9] = 2$$

$$[2.1] = 2$$

$$[-1] = -1$$

$$[-1.1] = -2$$

$$[-1.9] = -2$$