

INEQUALITY SOLVING

MATH 152, SECTION 55 (VIPUL NAIK)

Difficulty level: Moderate for people who have been inequalities before, high for people who haven't.

Corresponding material in the book: Section 1.3. However, this section does *not* cover inequalities involving rational functions, and only does the polynomial case. The rational functions case is only a mild generalization of the polynomial case, *but it is essential that you understand the rational functions case from the notes.*

Things that students should definitely get: Set notation for writing intervals, how to manipulate and solve inequalities, inequalities for rational functions, inequalities for absolute value.

EXECUTIVE SUMMARY

Words...

- (1) When solving inequalities, we can add the same thing to both sides. We can subtract the same thing from both sides. We can add two inequalities with the same direction of inequality.
- (2) We can multiply a positive number to both sides and preserve the direction of inequality. If we multiply by a negative number, we reverse the direction of inequality. If we multiply by a number that we know is nonnegative (But we're not sure if it is positive or zero) then the direction of inequality is preserved but an equality case gets introduced. e.g., if $a > b$ and $x \geq 0$, we get $ax \geq ab$.

Actions (think back to examples where you've dealt with these issues)...

- (1) If $|x|$ takes values in a certain set A , then the set of possible values for x is the union of A and the negatives of numbers in A .
- (2) Generally, when solving inequalities involving absolute value, consider the case that the thing inside the absolute value is negative and the case that it is nonnegative. Solve both cases and then take the union of the solutions.
- (3) When solving inequalities involving rational functions, bring everything to one side. So it reduces to trying to find when a given rational function is positive, negative, zero, and undefined.
- (4) A rational function is undefined wherever the denominator is 0. It is 0 where the numerator is 0 but the denominator isn't. For positive and negative, start from the far positive end and remember that every time you cross over a linear factor of the numerator or the denominator, the sign changes...
- (5) ... except that the sign changes only if the total multiplicity of that factor is odd. If the total multiplicity is even, then the sign doesn't change. (What do I mean? Review the examples we did and try figuring out)...
- (6) ... and what if there are quadratic factors in the numerator or denominator, such as $x^2 + 1$, that do not factorize further? These are anyway always positive or always negative, so they don't affect the sign of things ...
- (7) ... and what if there are quadratic factors that you don't know how to factorize? Well, use the quadratic formula.

1. EQUALITY VERSUS INEQUALITY

In this lecture, we cover inequalities. This is almost like solving equations, except that instead of an equality sign, there is an inequality sign. The main difference is that because we have an inequality sign, the solution sets are usually much bigger – they're not just a few isolated points, they are unions of intervals. Further, the solution to the corresponding *equation* is usually on the *boundary* of the solution set to the inequality.

2. REVIEW OF NOTATION

2.1. Review of set notation. So I hope that you've either seen the set notation before, either before coming here or in the last two days. But I'll review some of the basics anyway, very quickly. As you may have heard, a set is a *well-defined* collection of objects, called its *members* or *elements*. If A is a set and x is an object, we say $x \in A$ if x is a member of A , and $x \notin A$ otherwise. We have the symbols \cup for union of sets, \cap for intersection of sets. We have the symbol \emptyset for empty set, symbols \subset and \subseteq for strict and not necessarily strict containment. For reverse containment, we have \supset and \supseteq . And, as we saw last time, \setminus represents the set difference.

Now, let's discuss two ways of writing sets. One way is what we might call the *laundry list* or *naive* way: just list everything. I think a more professional word for this is the *roster method*. So, for instance, the set of all one-digit positive integers is $\{1, 2, 3, 4, 5, 6, 7, 8, 9\}$. That's a complete list of all of them.

The other approach, which is the *set builder* method or constructive method, specifies a qualification. For instance, if \mathbb{R} denotes the set of real numbers (and it always does in this course), then the set:

$$\{x \in \mathbb{R} \mid x > 0 \text{ and } x^2 + x + 1 > 2^x\}$$

is basically a set given by some condition. To determine whether an x is in the set, we check that it satisfies the condition. Is 1 in the set, for instance? Well, $1 > 0$, and $1^2 + 1 + 1 = 3$, which is greater than $2^1 = 2$. Good, so it is in. What about 0.5? Well, $0.5^2 + 0.5 + 1 = 1.75$, which is greater than $2^{0.5} = \sqrt{2}$.

Now the book does not use the \mid separator, it uses the $:$ separator, and that's fine too. So in the book's notation, the above becomes:

$$\{x \in \mathbb{R} : x > 0 \text{ and } x^2 + x + 1 > 2^x\}$$

2.2. Review of interval notation. Certain subsets of \mathbb{R} that come up pretty frequently are the *intervals*, and there is special notation for these. Let's discuss both the notation and the way these intervals are shown on the number line.

The interval (a, b) is the set $\{x \in \mathbb{R} : a < x < b\}$. In other words, it is the set of (real) numbers *strictly between* a and b . It is represented on the number line by creating unfilled circles at a and b and shading or darkening the region in between. This interval is termed the *open interval* between a and b . By the way, one of your homework problems asks you to think of this open interval in another way – in terms of a *center* and a *radius*. That's a very important homework problem not because it is particularly difficult, but because it is critical to many of the $\epsilon - \delta$ definitions of limits.

The interval (a, b) is also denoted $]a, b[$. The latter notation has both advantages and disadvantages. The primary advantage is that while (a, b) may be confused with an *ordered pair* representing a *point in the coordinate plane* with coordinates a and b , the notation $]a, b[$ has no alternative interpretations. However, for these notes and the rest of this course, we'll use the (a, b) -notation. This is used in the book and is also standard in most mathematics courses you will see.

Note, by the way, that the interval (a, b) is empty, and shouldn't be talked about, if $a \geq b$.

The interval $[a, b]$ is the set $\{x \in \mathbb{R} : a \leq x \leq b\}$. This is called the *closed interval* between a and b , and we use filled circles at a and b instead of the unfilled circles used earlier. Okay, here's a question: what is $[a, b] \setminus (a, b)$? In other words, what happens when you remove the open interval from the closed interval? What's the difference? The answer is: the two points a and b . So $[a, b] \setminus (a, b) = \{a, b\}$, and $(a, b) \cup \{a, b\} = [a, b]$.

There are also notions of half-open, half-closed intervals. So what does it mean for an interval to be left-open and right-closed? Well, that's an interval of the form $(a, b] = \{x \in \mathbb{R} : a < x \leq b\}$. And it's represented by an unfilled circle at a and a filled circle at b . In the other notation, it would be $]a, b]$. Similarly, $[a, b) = \{x \in \mathbb{R} : a \leq x < b\}$ is represented by a filled circle at a and an unfilled circle at b . In the other notation, it would be $[a, b[$.

So I think you're getting the general philosophy. The round parentheses $()$ represent openness, or the endpoint excluded, while the square braces $[]$ represent closedness, or the endpoint included. Pictorially, an excluded point is an unfilled circle, and an included endpoint is a filled circle.

To complete the discussion, we need to talk about ∞ . Now, ∞ is a big and mind-boggling concept and we're not really going to discuss it here. For our purposes, ∞ (positive infinity) is a placeholder for *no upper limit* and $-\infty$ is a placeholder for *no lower limit*.

So the interval (a, ∞) is the set $\{x \in \mathbb{R} : a < x\}$, and $[a, \infty)$ is the set $\{x \in \mathbb{R} : a \leq x\}$. And similarly, $(-\infty, a)$ is the set $\{x \in \mathbb{R} : x < a\}$ and $(-\infty, a]$ is the set $\{x \in \mathbb{R} : x \leq a\}$. And on the number line, you just use an arrow to indicate that it'll go on forever.

Notice that the parentheses around ∞ are always the round ones, meaning that ∞ is never included. And that makes sense because ∞ isn't real. Nor is $-\infty$. Whether these things exist and what they mean is beyond the scope of our discussion. I just want you to think of them as placeholders.

2.3. As a union of intervals. Okay, here's a quick test. Express $(-1, 1) \setminus \{0\}$ as a union of intervals? What is it? [Draw diagram]. You see the picture. It is $(-1, 0) \cup (0, 1)$.

Or, what about $(-2, 2) \setminus (-1, 1)$? [Draw diagram]. That's $(-2, -1] \cup [1, 2)$. Note the way the circles fill and unfill.

3. INEQUALITY SOLVING

3.1. Some basic rules. Let's discuss some very basic rules of inequality-solving. The way you solve inequalities is very similar to the way you solve equalities, except this: when you multiply both sides by a negative number, you change the direction of the inequality. And the one thing you shouldn't do is multiply both sides by zero.

Okay, so consider the inequality:

$$x + 4 \leq 5(x - 1)$$

Moving all stuff to one side gives:

$$-4x + 9 \leq 0$$

Now, you multiply both sides by -1 , so change the sign of the inequality, and get:

$$4x - 9 \geq 0$$

And then simplify to $x \geq 9/4$. Which is the interval $[9/4, \infty)$ on the number line.

Inequality-solving for polynomial functions. Let's now consider some polynomial functions. The goal is to consider a polynomial function f , and try to determine where it is zero, where it is positive, and where it is negative. For now, we'll focus on polynomial functions that split completely into linear factors.

For instance, consider the polynomial function:

$$f(x) = x(x + 1)(x - 1)$$

This polynomial is a product of three linear factors. How do we figure out the sign of this polynomial function at a point? Its sign is the product of the signs of the three factors. Now, if you have a linear factor $x - a$, where is it positive, where is it zero, and where is it negative? Answer: it is positive for $x > a$, zero for $x = a$, and negative for $x < a$.

So, we see that each linear factor switches sign at the corresponding zero (i.e., the linear factor $x - a$ switches sign at a). If only one linear factor switches sign, and the signs of the other factors remain the same, then the product also switches sign. So, for the above polynomial f , the sign changes could potentially occur at $-1, 0, 1$. Let's view this on a number line.

To the right of 1, x is pretty large, so all the three factors are positive. A product of three positives is a positive, so for $x > 1$, $f(x)$ is positive. At 1, the function becomes zero. Immediately to the left of 1, the factor $x - 1$ becomes negative, but the other two factors are still positive. So, the overall product is negative. And so the story goes till we hit $x = 0$. At $x = 0$, the function again takes the value 0. Then, to the immediate left of 0, both x and $x - 1$ are negative, but $x + 1$ is positive. So, the function is positive, and remains so till we reach $x = -1$. There it becomes 0, and to the left of -1 , all factors are negative, hence so is the product.

The upshot: $f(x) = 0$ for $x \in \{-1, 0, 1\}$, $f(x) > 0$ for $x \in (-1, 0) \cup (1, \infty)$, and $f(x) < 0$ for $x \in (-\infty, -1) \cup (0, 1)$.

Let's consider another function:

$$g(x) := x^2(x - 1)$$

The points where interesting things could happen are, in this case, 0 and 1. But 0 is *doubly* interesting, because the linear factor x has multiplicity 2.

So, to the right of 1, all factors are positive, so $g(x) > 0$ for $x > 1$. At 1, g takes the value 0. To the immediate left of 1, $x - 1$ becomes negative but the remaining factors are still positive. So the function becomes negative. Then, at 0, it becomes 0. What happens to the left of 0? Does the function switch sign again? No! And that's because although the x switches sign to negative, there are two of them, so their effects cancel each other. So the overall effect is no sign change, and the function remains negative.

So $g(x)$ is positive for $x \in (1, \infty)$, it is negative for $x \in (-\infty, 0) \cup (0, 1)$, and it is zero for $x \in \{0, 1\}$.

3.2. Inequality solving for rational functions. *This is not described explicitly in the book, but is important for some later material, so please go through it carefully. We'll try to briefly go over it in lecture.*

Let's now consider the rational function $P(x)/Q(x)$, where P and Q are polynomials. We want to figure out where this function is positive, zero, and negative.

Now the first thing you should remember about rational functions is that they aren't always globally defined. The points where they aren't defined are the points where the denominator implodes and the expression explodes: the roots of Q . So at these points, there is no inequality satisfied: the function just isn't defined.

The first thing we do is factor both P and Q . So we obtain:

$$\frac{a(x - \alpha_1)(x - \alpha_2) \dots (x - \alpha_m)}{b(x - \beta_1)(x - \beta_2) \dots (x - \beta_n)}$$

So, we already figured out that the values $\beta_1, \beta_2, \dots, \beta_n$ are points where this function isn't defined. And by the way, do not cancel before figuring out these points, because we have to take the function and treat it *as is* (remember FORGET?). But once we've excluded all these points, then away from these points, we can cancel, so let's assume from now on that there is no common factor between the numerator and the denominator.

Also, a/b has some sign (positive or negative) so let's ignore that too, because that'll just flip signs for everything. So we're really looking at:

$$\frac{(x - \alpha_1)(x - \alpha_2) \dots (x - \alpha_m)}{(x - \beta_1)(x - \beta_2) \dots (x - \beta_n)}$$

The first thing to observe is that the overall sign is the product of signs of each of the pieces. And the sign of $x - t$ is positive if x is greater than t and negative if x is less than t . So each time x crosses one of the α_i s or β_j s, one expression flips sign. Thus, the only points where the expression changes sign, or crosses zero, are the α_i s and β_j s [See caveat to this in a later example].

Okay, now if your head is already swimming with this, it's time to take a simple example. Now if I were trying to impress you with my pedagogy, I would probably start with the simplest example, then gradually build up to more and more complex examples. And that's a very valuable approach that I'll use most of the time. But, often, I'll start by just trying to attack the general problem, make a bit of headway, get stuck, and *then* proceed to a simple example. It's important that you develop a tolerance for both styles because you'll need both styles. There are times when building up gradually from one example to another isn't a luxury you can afford, and there are times where it's the only thing that makes sense.

So let's consider:

$$\frac{x(x + 1)(x + 2)}{x(x - 1)(x - 2)}$$

So what are the points where this function isn't defined? 0, 1 and 2, the three roots of the denominator. Note that at 0, our worry about undefinedness is a little silly because if we were a little smarter we'd cancel

x and get a new function that is defined at 0. But as written, the function is undefined at 0. [In the limit/continuity jargon, the function has a *removable* discontinuity at 0.]

Anyway, away from these three points, the function simplifies to:

$$\frac{(x+1)(x+2)}{(x-1)(x-2)}$$

So, the points where we could have a sign change are $-2, -1, 1, 2$. Let's start from the right. On the far right, everything is positive, because x is greater than all four numbers. So the expression is positive. When you cross 2, $x-2$ becomes negative, the others remain positive. So, negative. Then, when you cross 1, both $x-1$ and $x-2$ become negative, the others remain positive, so positive. And then after -1 it becomes negative and after -2 again positive. So the function is positive at $(-\infty, -2) \cup (-1, 1) \cup (2, \infty)$. But wait. The *original* function wasn't defined at 0, so we need to exclude that. So the original function is positive on $(-\infty, -2) \cup (-1, 0) \cup (0, 1) \cup (2, \infty)$. The function is negative at $(-2, -1) \cup (1, 2)$. The function is zero on $\{-2, -1\}$. It is undefined at $\{0, 1, 2\}$. By the way, there's an important difference between the point 0, where it is undefined only because we wrote the function stupidly, and $\{1, 2\}$, where it is undefined for more fundamental reasons. Namely, at 0, the discontinuity is removable, but at 1 and 2, it cannot be removed.

So we have a pretty good feel. But there's an important thing that we didn't take into account: higher powers.

For instance, consider:

$$\frac{x(x+1)^3}{(x-1)^2}$$

So here, the function is not defined at 1. It's positive to the right of 1. What happens as we cross 1? The $(x-1)$ term changes sign, but it is squared, so the sign change doesn't affect the sign of the overall expression. And indeed, there's no sign change. So the expression continues to be positive. Then, at 0, there is a sign change, and the expression becomes negative. At -1 , there is another sign change, and the expression becomes positive again. The reason? The exponent in this case is odd.

So, the function is positive on $(-\infty, -1) \cup (0, 1) \cup (1, \infty)$, negative on $(-1, 0)$, zero on $\{-1, 0\}$, and undefined at 1.

To view a consolidated summary on how to handle inequalities involving rational functions, go to the executive summary at the beginning of the document.

4. INEQUALITIES AND ABSOLUTE VALUE

4.1. Absolute value. We've talked about the absolute value function, so I'll just say a bit about solving inequalities involving the absolute value function. The main thing to remember is: if the absolute value of a number is in a set A of nonnegative numbers, then that number itself is in the set $A \cup -A$, where $-A$ is the set of negatives of elements of A . So, for instance:

$$|x-2| < 3$$

is equivalent to:

$$-3 < x-2 < 3$$

So an inequality involving absolute values is really an inequality that involves both an upper bound and a lower bound. So it is two inequalities rolled into one. Now, when you have two inequalities rolled into one, you can solve them separately and *intersect* the solutions. But in this case, it's easy to sort of solve them both together. Just add 2 to both sides:

$$-1 < x < 5$$

so the solution set is the open interval $(-1, 5)$.

I suggest you look at more examples in the book involving absolute values, rational functions and inequalities. There are a lot of tricks. These ideas are very important. They'll come up again and again, particularly in the context of $\epsilon - \delta$ proofs.

4.2. More ways of thinking about absolute values. Absolute values and the inequalities related to them could be a little tricky, so here are some further intuitive ways of thinking about the absolute value function that might be useful when thinking about inequalities.

One way of thinking of the absolute value function is as a folding function. What I mean is, you think of the number line as a long thin strip of paper, and to calculate the absolute value, you fold it about 0, so that $-x$ comes on top of x . So *undoing* the absolute value function is like *unfolding*.

Let's take some examples. Suppose you know that the absolute value of x is 3. Then, what can you say about x ? You may say that $x \in \{-3, 3\}$. In words, $x = 3$ or $x = -3$. A convenient shorthand is $x = \pm 3$. You may have seen this kind of shorthand in the formula for the roots of a quadratic equation.

Now, suppose you are given that the absolute value of x is either 2 or 3. What are the possible values of x ? Think again about unfolding, and you'll see that the possibilities for x are ± 2 and ± 3 .

Okay, by the way, what can you say about x if you are given that $|x|$ is either 2 or -1 ? The thing you have to say is that the -1 case is always false. It has no solution. So the only legitimate case is $|x| = 2$, which gives $x = \pm 2$.

Now we can proceed to thinking about inequalities. Suppose we are given that $2 < |x| < 3$. What does this tell you about the possibilities for x ? Unfolding, you see that it's either the case that $2 < x < 3$ (that happens when x is positive) or $-3 < x < -2$ (That happens when x is negative). So the solution set is $(-3, -2) \cup (2, 3)$.

Okay, what about $-1 < |x| < 2$? Well, the first thing you should note is that $-1 < |x|$ is a zero information statement, because it's always true. [SIDENOTE: A statement that is always true is termed a *tautology* and a statement that is always false is termed a *fallacy*.] So we can discard that, and we get $|x| < 2$. So, $|x|$ is in the region $[0, 2)$. Unfolding this, we obtain that the possibilities for x are the interval $(-2, 2)$.

What about $0 < |x| < 2$? Well, you can guess the answer by now: it is $(-2, 0) \cup (0, 2)$.