Math 314 Class Notes, Spring 2024 (being revised live from Fall 2019)

Dr. Holmes

updated for spring 2024 class: correction on 1/8/2024 of a typo noticed by a student.
updated Jan 21 to fix error in date of first homework assignment.

Contents

Documents of this kind *always* include horrible typos and mental slips. Please call these to my attention when you find them or suspect you have found them. I give **bonus points** for finding typos (to the first person who finds them other than myself).

To my Spring 2024 students: For the moment, I have cloned the Fall 2019 notes, but you should expect changes. Read the section A Sneak Preview to review the intended first lecture, and do the assignment which I am inserting into the notes right after that section (when I actually tell you on the class announcement page that it is assigned). You will have already seen it on the board. This document contains lots of assignments with due dates in 2019 (I suppose it might contain ones with even earlier due dates!) Nothing is assigned to you yet until it has a Spring 2024 date on it, and assignments will be announced separately on the class announcements page.

1 Intentions of this Course

This course is titled Introduction to Real Analysis. Real Analysis is the study of the basic properties of the real number system: from your standpoint, the purpose of real analysis is to provide a foundation for the calculus you learned in Math 170 and Math 175. The concept of limit is introduced intuitively in Math 170 (Calculus I, if you are a transfer) [limit of a function; limit of a sequence is introduced in Math 175 and the same comments apply): while a formal definition of the notion(s) of limit is given, we really do not work with it, instead giving theorems to allow us to work with limits which are "intuitively obvious" but which we never prove. We will prove them in this class. Two very important theorems, the Intermediate Value Theorem and the Extreme Value Theorem, appear early in Math 170 and are used extensively, but are quite impossible to prove given what we know in Math 170. Proving these theorems is an essential part of my agenda in this course.

There are other things that we pass over in calculus which I would like to do in this class if we have time (but we very likely will not, unless I choose to take the course in a different direction this time). The definition of the integral given in Math 170 books is often quite dishonest; certainly the existence of the integral of a continuous function is not really proved. I would love to investigate this, though historically I usually do not get there. The definition of the derivative presents no difficulties (once the definition of limit is thoroughly explored), but it might be interesting to properly investigate

the proofs of such basic results as the Chain Rule, the Mean Value Theorem, and L'Hôpital's Rule. However, it should be noted that I did not even define derivative or integral in this class last term: we spent the entire course working on things mentioned in the previous paragraph.

There is not a fixed body of real analysis content that we need to get to in this course: I have discussed this with the M414 instructors, and they understand that what gets covered in M314 is variable, and they are likely to cover the basics (more quickly) from the beginning.

A crucially important aspect of this course has to do with method not content. This is a proofs course. You will read and write proofs in this course. Because of the importance of proofs in this course, we will begin by doing propositional logic and logic of quantifiers in the abstract, and then do proofs in formal arithmetic (you will get a real workout in mathematical induction!), and only then get to real analysis. You will get some exposure to my own research in the logic part of the course: we will do some labs on formal proofs in logic using my research software. The logic content is not covered in the textbook (which in fact uses little or no logical symbolism), thus I am writing these notes. I will very likely continue to write some notes of my own when we do get into the text.

2   A Sneak Preview

I would like you to understand why we have to look carefully at logic and proof strategy in this course. The best way to do this is to start out by proving some things that we "should have" proved in Math 170. This will show you that we have to be much more careful about how we read and write mathematics in this course, and help to justify our departure for a time to the realm of logic.

We begin with a definition we should know from calculus, though something important is often left out of it.

Definition:
$$\lim_{x \to a} f(x) = L$$

means "For each $\epsilon > 0$, there is a $\delta > 0$ such that for any x, if $0 < |x - a| < \delta$, then $|f(x) - L| < \epsilon$".

That is quite a mouthful. The phrase "for any x" is often left out.

A basic thing to notice, which is essential to understanding what this definition says, is that for any real numbers a, b, the quantity $|a - b|$ is the *distance* from a to b. So $0 < |x - a| < \delta$ says that x and a are distinct and within distance δ of one another, for example.

We prove a statement of a sort that you very likely did prove, or see proved, in calculus.

Theorem:
$$\lim_{x \to 3} 10 - 2x = 4$$

What this means: Applying the definition, this means "For any $\epsilon > 0$, there is a $\delta > 0$ such that for any x, if $0 < |x - 3| < \delta$ then $|(10 - 2x) - 4| < \epsilon$".

Scratch work: The overall strategy will be to suppose that we are given a positive real number ϵ about which we know nothing, and figure out what δ has to be to make this statement true. The reason that this paragraph is titled "scratch work" is that we are going to work backwards to figure out what δ has to be, then write the proof in the forward direction when we have worked this out. What we want as our final conclusion is $|(10 - 2x) - 4| < \epsilon$: this is supposed to follow from $0 < |x - 3| < \delta$. $|(10 - 2x) - 4| = |6 - 2x| = 2|3 - x| = 2|x - 3|$. Thus $|(10 - 2x) - 4| < \epsilon$ is equivalent to $2|x - 3| < \epsilon$ and so is also equivalent to $|x - 3| < \frac{\epsilon}{2}$. This suggests that we should take δ to be $\frac{\epsilon}{2}$.

Proof: Let ϵ be an arbitrarily chosen positive real number. Let $\delta = \frac{\epsilon}{2}$. Let x be chosen so that $0 < |x - 3| < \delta = \frac{\epsilon}{2}$. Then $2|x - 3| < \epsilon$, so we have $|(10 - 2x) - 4| = |6 - 2x| = 2|3 - x| = 2|x - 3| < \epsilon$. We have proved the theorem.

While this was shown to you in Math 170, and you may have learned to write a proof of this kind in a stereotyped manner to answer an exam question, I am sure that many of you did not understand it, and moreover it was not necessary to understand this fully to understand enough about limits to get on in Calculus I.

Now I will present a second theorem of a sort you probably did not see in Math 170, though there *are* instructors who might present a proof of this kind in class in Calculus I, though very few who would expect you to write such a proof on an exam at that level. The statement proved is perfectly well-known to you and you even know how to write a sort of proof of it.

Theorem: $\lim_{x \to 3} x^2 = 9$

What this means: For each $\epsilon > 0$, there is a $\delta > 0$ such that for any x, if $0 < |x - 3| < \delta$, then $|x^2 - 9| < \epsilon$.

Scratch work: Our strategy for proving this is going to be to choose a positive real number ϵ arbitrarily then find a δ depending on that ϵ for which $|x^2 - 9| < \epsilon$ for any x with $0 < |x - 3| < \delta$. As in the previous proof, we work backward from the conclusion $|x^2 - 9| < \epsilon$. We factor, getting $|x - 3||x + 3| < \epsilon$, then $|x - 3| < \frac{\epsilon}{|x+3|}$. We cannot just set $\delta = \frac{\epsilon}{|x+3|}$, because δ can depend on the choice of ϵ, but not on the choice of x. We can make assumptions about δ which will help, though. Notice that if any δ works, so does any smaller one. So it is harmless to stipulate $\delta \leq 1$. This means that if $|x - 3| < \delta$, we have $|x - 3| < 1$, so $-1 < x - 3 < 1$, so $2 < x < 4$, so $5 < x + 3 < 7$. We have $|x + 3| = x + 3$ because $x + 3$ is positive. Now $x + 3 = |x + 3| < 7$ is also equivalent to $\frac{\epsilon}{7} < \frac{\epsilon}{|x+3|}$, and from this we see that $\delta = \min(1, \frac{\epsilon}{7})$ should work.

Proof: Choose an $\epsilon > 0$. Set $\delta = \min(1, \frac{\epsilon}{7})$. Notice that $\delta \leq 1$ and $\delta \leq \frac{\epsilon}{7}$. Now choose an arbitrary x such that $0 < |x - 3| < \delta = \min(1, \frac{\epsilon}{7})$. We have $|x - 3| < 1$, so we have $5 < |x + 3| < 7$ by the argument given above in the scratch work. Notice that we also have $|x - 3| < \frac{\epsilon}{7}$. Now $|x^2 - 9| = |x + 3||x - 3| \leq 7|x - 3| < 7\frac{\epsilon}{7} = \epsilon$, which completes the proof.

Now we prove a more general theorem, one of the basic limit theorems that we are given in Math 170 but are never asked to prove. Actually, I have proved this particular basic limit theorem in class in Math 170, but I have never expected students to be able to reproduce the proof. In this class, you will be expected to be able to produce such proofs.

Triangle Inequality: Recall that $|a - b|$ is the distance between two real numbers a and b. The geometric Triangle Inequality says that $d(A, C) \leq d(A, B) + d(B, C)$ for any A, B, C. We do this in the context of the real number line with $A = x; B = 0; C = -y$. We get $|x - (-y)| \leq |x - 0| + |0 - (-y)|$, that is $|x + y| \leq |x| + |y|$. I am sure you have seen this second identity, and seen it called the Triangle Inequality, but I wonder if you have ever seen an explanation of its relation to the geometric identity.

Theorem: If $\lim_{x \to a} f(x) = L$ and $\lim_{x \to a} g(x) = M$, then

$$\lim_{x \to a} [f(x) + g(x)] = L + M.$$

Scratch work: Choose an $\epsilon_0 > 0$ arbitrarily (you will see a reason why we want a name for this which is not simply ϵ, but it is still completely arbitrarily chosen). Our final aim is to get $|[f(x)+g(x)]-[L+M]| < \epsilon_0$. We note that $|[f(x) + g(x)] - [L + M]| = |(f(x) - L) + (g(x) - M)| \leq |f(x)-L|+|g(x)-M|$. If we make each of $|f(x)-L|$ and $|g(x)-M|$ less than $\frac{\epsilon_0}{2}$ then their sum will be less than ϵ_0. Now because $\lim_{x \to a} f(x) = L$, we can find δ_1 such that for any x such that $0 < |x - a| < \delta_1$, we have $|f(x) - L| < \frac{\epsilon_0}{2}$ and because $\lim_{x \to a} g(x) = M$, we can find δ_2 such that for any x such that $0 < |x - a| < \delta_2$, we have $|g(x) - M| < \frac{\epsilon_0}{2}$. So we want $\delta = \min(\delta_1, \delta_2)$.

Proof: Suppose that $\lim_{x \to a} f(x) = L$ and $\lim_{x \to a} g(x) = M$. Choose $\epsilon_0 > 0$ arbitrarily. Choose a δ_1 such that for any x such that $0 < |x - a| < \delta_1$, we have $|f(x) - L| < \frac{\epsilon_0}{2}$ (we can do this because $\lim_{x \to a} f(x) = L$). Choose a δ_2 such that for any x such that $0 < |x - a| < \delta_2$, we have $|g(x) - M| < \frac{\epsilon_0}{2}$ (we can do this because $\lim_{x \to a} g(x) = M$). Let $\delta = \min(\delta_1, \delta_2)$. Choose x so that $0 < |x - a| < \delta$, but otherwise arbitrarily. Notice that $0 < |x - a| < \delta_1$ and $0 < |x - a| < \delta_2$, so $|f(x) - L| < \frac{\epsilon_0}{2}$ and $|g(x) - M| < \frac{\epsilon_0}{2}$, so $|f(x) - L| + |g(x) - M| < \epsilon_0$. Then $|[f(x) + g(x)] - [L + M]| = |(f(x) - L) + (g(x) - M)| \leq$ [triangle inequality] $|f(x) - L| + |g(x) - M| < \epsilon_0$. This completes the proof.

If you find these proofs forbidding, you should begin to get an idea why we are going to take a detour through logic and proof strategy. The content subtleties here about absolute values are not what makes it hard. You are likely not to be as well prepared for reasoning about inequalities as for reasoning about equations, but it is really the logical reasoning aspect that makes proof hard.

assignment on the next page

2.1 Assignment due Wed, Jan 24, 2024

Prove the assertion
$$\lim_{x\to 1}(3x + 2) = 5$$
. Use the style of the first example above. Be sure to start by expanding the definition of limit, that is, writing out in full what this statement means before starting to prove it. This is due at the beginning of class on Wed Jan 22.

Further, expand the statement $\lim_{x\to a}(f(x)g(x)) = LM$ which might be the conclusion of a statement of the multiplication property of limits, by expanding the definition of limit. Your statement should mention only logical and precalculus concepts. A model for the expansion I have in mind is found in the whiteboard notes from the Wed Jan 17th lecture (look at what I do with the sum rule for limits).

I remind you that homework is at this point to be turned in to me by email. LaTeX or scanned handwritten documents are acceptable. Please be certain that the name of your file that you turn in includes your last name, identification of the assignment, and identification of the course.

You are always encouraged to see me in my office if you have questions!

3   Propositional Logic

[NOTE: I started here by lifting a block of text straight from last term's notes (which you can also look at); I'm going to add some additional remarks based on things I said in class this time.]

This section is about the logic of little words like "not", "and", "or", "if". Mathematicians use these words in particular ways. We will introduce symbols for the concepts usually represented by these words; one should remember that everything I do with logical symbolism is intended to convey something about formal proofs written in mathematical English. You are welcome to use the formal symbolism in proofs later in the course, if you find that it helps you to express yourself clearly, but it will seldom if ever be required. The points I make about proof strategy using the symbolism will be needed all the time, though.

3.1   A menagerie of logical operations

Each of these operators will be defined using truth tables. This does not mean that we will make extensive use of the method of truth tables, which is a complete proof technique for propositional logic but a very inefficient one and one which does nothing to help us see how to write proofs. The advantage of truth tables is that they are very clear.

Each logical operator has its own Latin name; it's a good idea to know the Latin names just because I use them...

Capital letters $P, Q, R \ldots$ are used here to stand for simple statements – all we care about about these simple statements is whether they are true or false.

3.1.1   Not (negation)

If P is a statement, we denote "It is not the case that P" (often abbreviated "not P" or "P is false", but try substituting an English sentence for P in either of these...) by the notation $\neg P$.

P	$\neg P$
T	F
F	T

3.1.2 And (conjunction)

If P is a statement and Q is a statement, we denote "P and Q" by the notation $P \wedge Q$.

P	Q	$P \wedge Q$
T	T	T
T	F	F
F	T	F
F	F	F

The English word "and" when it connects sentences generally means the same thing as $\wedge$. When it connects predicates as in "21 is composite and divisible by 7" it abbreviates a conjunction of sentences "21 is composite and 21 is divisible by 7". It is all right to write "21 is composite $\wedge$ 21 is divisible by 7"; we do not write "21 is composite $\wedge$ divisible by 7". We only use $\wedge$ to connect sentences. When the English word "and" connects noun phrases it sometimes abbreviates a use of $\wedge$, as in "John and Mary like ice cream", which could be written "John likes ice cream $\wedge$ Mary likes ice cream", but sometimes it doesn't have anything to do with logical conjunction, as in "John and Mary moved the half-ton safe" or "21 and 55 are relatively prime". Notice that uses of "and" of this latter kind do happen in mathematics.

3.1.3 Or (disjunction)

If P is a statement and Q is a statement, we denote "P or Q" by the notation $P \vee Q$

P	Q	$P \vee Q$
T	T	T
T	F	T
F	T	T
F	F	F

The table here actually tells us that we are going to use "or" in a more specific way than English requires. English usage allows both "inclusive or", which lawyers write "and/or" and "exclusive or". The sense of "or" that is always used in mathematics (at least this is our official position) is the inclusive sense, "P or Q or both". We provide a symbol $P \oplus Q$ with a table

P	Q	$P \oplus Q$
T	T	F
T	F	T
F	T	T
F	F	F

for the exclusive sense, which computer scientists call "xor" and which occurs in the sentence "You will have chocolate ice cream or you will have vanilla ice cream" uttered in a certain maternal tone. We will seldom if ever actually use this logical construction.

Remarks about the use of "or" between verb phrases or noun phrases in English are similar to those about "and", except that such English uses generally do abbreviate a use of $\vee$ (or of $\oplus$).

3.1.4 If...then... (implication, conditional)

If P is a statement and Q is a statement, we define $P \rightarrow Q$ using the table

P	Q	$P \rightarrow Q$
T	T	T
T	F	F
F	T	T
F	F	T

and we claim that this is what we mean when we say "If P, then Q" in mathematical English. The only bit of this that we genuinely agree with as native speakers of English is that "If P, then Q" is false if P is true and Q is false. "If Vladimir Putin is Prime Minister of Russia, then Napoleon Buonaparte was Emperor of France" is true according to the table and causes us to scratch our heads as English speakers, because we really would like to see a relationship between the two statements. "If Napoleon conquered China then $2 + 2 = 5$" is true according to the table but seems very doubtful: after all, in a perhaps unlikely alternate history where Napoleon *did* conquer China, we still suspect that two plus two would be four, even if the Emperor decreed otherwise. The point here is that the mathematical sense of implication has nothing to do with causality or counterfactual speculations.

A mathematician is content to say that anything implies a true statement, and a false statement implies anything, and these two considerations are enough to fill in the table above. We admit absolutely that this is not the

12

conventional use of "if. . .then. . .", and it might not always even be the correct translation of uses of this phrase in your math book. It will however be the translation of uses of this phrase in English statements of theorems in your math book.

Other ways of saying $P \to Q$: P only if Q, P implies Q, P is sufficient for Q.

Other ways of saying $Q \to P$: P if Q, P is necessary for Q.

3.1.5   if and only if; iff (biconditional, equivalence)

If P is a statement and Q is a statement, the elaborate statement "P if and only if Q" is represented by the symbolic expression $P \leftrightarrow Q$ and is accurately captured by the following table. A common abbreviation for this used in mathematical English is "P iff Q.

P	Q	$P \leftrightarrow Q$
T	T	T
T	F	F
F	T	F
F	F	T

This is equivalent to $(P \to Q) \wedge (Q \to P)$. It is interesting but not important that it is also equivalent to $\neg(P \oplus Q)$.

3.1.6   the absurd, contradiction

We introduce the symbol $\bot$ for a fixed false statement.

3.2   Complicated expressions

We officially adopt an order of operations: when there are no parentheses, we first apply $\neg$ (so a negation applies just to the immediately following letter or expression in parentheses), then apply $\wedge$, then apply $\vee$ [by which we mean that $P \wedge Q \vee R$ means $(P \wedge Q) \vee R$], then apply $\to$, then apply $\leftrightarrow$ (or $\oplus$ if we ever use it). All the binary operators except $\to$ are associative (and mixed expressions in $\leftrightarrow$ and $\oplus$ actually have the same meaning no matter how parentheses are placed): we interpret $P \to Q \to R$ as $P \to (Q \to R)$ (but we will usually write the parentheses). We "officially" adopt this: we will generally avoid relying on it, except that we will always regard $\neg$ as

applying just to the immediately following letter unless parentheses force it to apply to a larger expression.

Some useful equivalences: notice that $\neg P \vee Q$ is equivalent to $P \rightarrow Q$, and $\neg P \rightarrow Q$ is equivalent to $P \vee Q$.

Complex expressions involving negation are very awkward to express in English directly: in fact, we usually perform logical transformations on them rather than say them. "It is not the case that roses are red and violets are green" would be more likely to be expressed as "Roses are not red or violets are not green" (it does *not* mean "Roses are not red and violets are not green" and "It is not the case that snow is green or milk is blue" means "Snow is not green and milk is not blue" (not "Snow is not green or milk is not blue"). These are examples of de Morgan's Laws, which we note here but do not yet officially adopt as logical rules: $\neg(A \wedge B)$ is equivalent to $\neg A \vee \neg B$ and $\neg(A \vee B)$ is equivalent to $\neg A \wedge \neg B$. We can also observe that $\neg(P \rightarrow Q)$ is equivalent to $P \wedge \neg Q$, and we already noticed that $\neg(P \leftrightarrow Q)$ is equivalent to $P \oplus Q$ and $\neg(P \oplus Q)$ is equivalent to $P \leftrightarrow Q$. Of course $\neg\neg A$ is equivalent to A. I repeat that we notice all these equivalences as possibly important for reading complicated statements correctly but we have not yet adopted them as logical rules.

3.3 Truth tables

In this section we describe a complete method of proof for propositional logic which is interesting because it is complete but not really useful in practice.

If we have a sentence in the language of propositional logic, containing n letters, there are 2^n ways to assign truth values to the letters in the sentence. We can evaluate each of these and if all 2^n of them come out true, then the sentence is a theorem. A theorem of this kind is called a *tautology*. We can also determine when two sentences are essentially the same: if they have the same truth value for each of the 2^n assignments of truth values to the n letters appearing in both, then they are "the same" and are called *logically equivalent*.

Theorem:
$$((P \wedge Q) \rightarrow R) \leftrightarrow (P \rightarrow (Q \rightarrow R))$$

Comments: There are three variables, so there are $2^3 = 8$ combinations of true values to consider. We present them in a convenient table. The

rule for the construction of the table is that the values of each expression in the theorem appear in the column under its main operator. The column under any letter is a copy of the column under the same letter to the far right. So for example the column under the $\wedge$ is constructed by applying the "and" operation to the values under the P and Q in the same row. Then the value under the first $\rightarrow$ is computed by applying the "implies" operation to the truth values under the $\wedge$ (values of $P \wedge Q$) and the values under the R, in the same row, and so forth.

Truth Table:

P	Q	R	$((P$	$\wedge$	$Q)$	$\rightarrow$	$R)$	$\leftrightarrow$	$(P$	$\rightarrow$	$(Q$	$\rightarrow$	$R))$
T	T	T	T	T	T	T	T	T	T	T	T	T	T
T	T	F	T	T	T	F	F	T	T	F	T	F	F
T	F	T	T	F	F	T	T	T	T	T	F	T	T
T	F	F	T	F	F	T	F	T	T	T	F	T	F
F	T	T	F	F	T	T	T	T	F	T	T	T	T
F	T	F	F	F	T	T	F	T	F	T	T	F	F
F	F	T	F	F	F	T	T	T	F	T	F	T	T
F	F	F	F	F	F	T	F	T	F	T	F	T	F

Closing comments: Since the values in the column under the $\leftrightarrow$ are all T, the whole biconditional expression is a theorem (a tautology). This table also demonstrates that $(P \wedge Q) \rightarrow R$ and $P \rightarrow (Q \rightarrow R)$ are logically equivalent. This is a general phenomenon: if we can prove that $A \leftrightarrow B$ is a tautology (where A and B are complicated expressions) then we have also proved that A and B are logically equivalent.

The example should make it easy for us to see why the method of truth tables is not practical. The sentence

$$((P \rightarrow Q) \wedge (Q \rightarrow R) \wedge (R \rightarrow S) \wedge (S \rightarrow T)) \rightarrow (P \rightarrow T)$$

is pretty obviously a theorem and admits a short proof using the proof strategies we will introduce below. But there are 5 variables in it. so a truth table proof of the theorem would have 32 lines, each with quite a number of T's and F's.

3.4   Reduction of Operations

We first indicate using an example why the operations $\wedge$, $\vee$, and $\neg$ (AND, OR and NOT) are enough to represent any expression with a truth table.

For example in the table

P	Q	R	???
T	T	T	F
T	T	F	T
T	F	T	F
T	F	F	T
F	T	T	F
F	T	F	T
F	F	T	F
F	F	F	T

we can fill in ??? as

$$(P \wedge Q \wedge \neg R) \vee (P \wedge \neg Q \wedge \neg R) \vee (\neg P \wedge Q \wedge \neg R) \wedge (\neg P \wedge \neg Q \wedge \neg R).$$

The rule for construction of this expression is that it is a disjunction of one expression for each line of the truth table with a T in it. The expression corresponding to a given line with a T in it is the conjunction of expressions corresponding to the T's and F's in the row. The expression corresponding to a T in the row under a letter S is S; the letter corresponding to an F in a row under a letter S is $\neg S$. It is straightforward to see that this expression has the right truth table.

There is one exceptional case: if there is no T under the expression ??? then we can write it as $P \wedge \neg P$.

This general procedure shows that $\wedge$, $\vee$ and $\neg$ are in principle the only operations we need. Note that $P \to Q$ is equivalent to $\neg P \vee Q$ and $P \leftrightarrow Q$ is equivalent to $(P \wedge Q) \vee (\neg P \wedge \neg Q)$. In spite of this, we have practical reasons to want to use implication and biconditional.

In fact, we can reduce things further. We observe that $\neg (P \wedge Q)$ is equivalent to $\neg P \vee \neg Q$ (this is one of deMorgan's laws), so $P \wedge Q$ is equivalent to $\neg \neg (P \wedge Q)$ which is in turn equivalent to $\neg (\neg P \vee \neg Q)$.

This means that AND can be defined in terms of OR and NOT. [and in fact OR can be defined in terms of AND and NOT in basically the same way).

We can go even further. Define $P|Q$ (neither P nor Q) using the following truth table:

P	Q	$P|Q$
T	T	F
T	F	F
F	T	F
F	F	T

Notice that $P|Q$ is equivalent to $\neg(P \vee Q)$. The expression $P|P$ is then equivalent to $\neg(P \vee P)$ and so to $\neg P$ [it is easy to see that $P \vee P$ is equivalent to P], so we can define negation in terms of the neither/nor operator. Now $P \vee Q$ is equivalent to $\neg\neg(P \vee Q)$, that is $\neg(P|Q)$, that is $(P|Q)|(P|Q)$ [using the method for defining negation in terms of neither/nor that we just found]. So we can define both $\neg$ and $\vee$ in terms of $|$ (usually called NOR), and so we can define $\wedge$ in terms of $|$ (you can check that $(P|P)|(Q|Q)$ is equivalent to $P \wedge Q$), and so we can define anything at all with a truth table in terms of NOR!

There was a lot of excitement about this for some reason about 1900. Its interesting but not really useful; expressions in terms of NOR are not usually very efficient. To model the way that we actually reason in mathematics (and in fact the way we reason naturally in any context when we are being precise) we need all the operations we have listed. We don't gain by reducing the number of operations at the price of making logical sentences longer and harder to understand.

3.5 Care and Feeding of Implication

In this section, we discuss some special terminologies to do with the implication construction $P \rightarrow Q$.

"If P, then Q" means $P \rightarrow Q$.

"That P implies that Q" means $P \rightarrow Q$ (we often say "P implies Q" but it is not really correct English to say "Snow is white implies grass is green": it's really "That snow is white implies that grass is green". That said, I wrote "P implies Q" in class and likely will do so again.

"P if Q" means "$Q \rightarrow P$" (an example of how we must be careful).

The words "necessary" and "sufficient" are often used to signal implications.

"That x is odd is a necessary condition for x to be a prime greater than 2" is a true statement, and a fancy way of saying "x is a prime greater than 2" $\to$ "x is odd".

In summary "P is a necessary condition for Q" means $Q \to P$.

"That x is divisible by six is a sufficient condition for x to be divisible by two" is a true statement, and a fancy wasy of saying "x is divisible by 6" $\to$ "x is divisible by 2".

In summary, "P is a sufficient condition for Q" means $P \to Q$.

"P is a necessary and sufficient condition for Q" means $P \leftrightarrow Q$.

I am not myself a frequent user of the words "necessary" and "sufficient" in these senses, but they often appear in mathematical text, so you ought to know about them. And I might use them without even thinking about it.

The other terminology related to implication which I want to review is the use of the terms *converse*, *inverse* and *contrapositive*.

If we fix an implication $P \to Q$ that we are talking about, then its converse is $Q \to P$, its inverse is $\neg P \to \neg Q$, and its contrapositive is $\neg Q \to \neg P$.

If we believe the implication "If it is snowing, then it is below freezing at cloud level" then it is instructive to consider the sentences that result from this by these transformations:

original: If it is snowing, then it is below freezing at cloud level.

converse: If it is below freezing at cloud level, then it is snowing.

inverse: If it is not snowing, then it is not freezing at cloud level.

contrapositive: If it is not freezing at cloud level, then it is not snowing.

We do not believe the converse because we find it perfectly possible to imagine 70 degrees below zero and no humidity: without water, it won't snow. The inverse seems to make the same false claim that it will snow just because it is cold enough. The contrapositive, once we get our brains around it, seems to say the same thing as the original implication, but more indirectly.

Truth tables confirm these impressions.

P	Q	P	→	Q	Q	→	P	¬	P	→	¬	Q	¬	Q	→	¬	P
T	T	T	T	T	T	T	T	F	T	T	F	T	F	T	T	F	T
T	F	T	F	F	F	T	T	F	T	T	T	F	T	F	F	F	T
F	T	F	T	T	T	F	F	T	F	F	F	T	F	T	T	T	F
F	F	F	T	F	F	T	F	T	F	T	T	F	T	F	T	T	F

The columns giving the values of the whole expression are under the implication signs. Compare them and you can see that the original implication is always equivalent to its contrapositive, and the converse is equivalent to the inverse (which is the contrapositive of the converse, so its really the same fact).

3.6 Proof Strategies without Negation

We will present strategies for proving statements of each of the logical forms presented above. We will also present strategies for drawing further conclusions from statements of each of those forms which we have assumed or proved. These strategies will carry over to proofs written in mathematical English in the rest of the course.

In this subsection, we only discuss the strategies which do not involve negation (the logical operation $\neg$, "not").

We must carefully distinguish between two ways that a statement can appear in a proof. We may write a statement which we need to prove (or need to prove under current assumptions): we call this a "goal". We must be careful not to use a goal as if we could assume it to be true – do not assume what you are trying to prove. The other way that a statement can appear is if we have proved it, are assuming it for the sake of argument, or can show that it follows from local assumptions. There isn't a convenient word for this kind of statement: one might call it a "local conclusion" (where "local" reminds us that it may depend on assumptions we are making for the sake of argument in a local part of the proof.

Technical Remark about Letters: Letters from late in the alphabet (P, Q, R) are single propositions with no internal structure. Letters from earlier in the alphabet (A, B, C) stand for arbitrary sentences which might have complicated structure. P for example is not a theorem (because it can be either true or false) whereas A might be a theorem (it might be $P \vee \neg P$ on closer inspection).

3.6.1 Arguments, validity and soundness

This subsection introduces some terminology which will be used in talking about rules of reasoning.

An **argument** is defined as a pair of a collection of sentences (called "premises" or "hypotheses" of the argument) and a sentence called its "conclusion".

An argument is **valid** iff any assignment of values to variables appearing in the premises and conclusion which makes all the premises true also makes the conclusion true.

A sentence is a **theorem** iff it is the conclusion of a valid argument with the empty set of premises (so that any assignment of values to variables

appearing in the sentence makes it true).

An argument is **sound** iff it is valid and all the premises are true (independently of any assignment of values to variables). The conclusion of a sound argument is of course true.

An extended argument will often be presented as a list of numbered statements (in which some blocks of statements will be indented. Each line will be the conclusion of a valid argument with premises taken from earlier lines (and annotated in a way which makes it clear how it follows from which earlier lines). Sometimes an argument some of whose premises are assumed only for the sake of argument will be used: in such cases, the argument is indented as a block of lines, and lines in the block cannot be used outside the block (but the validity of the entire indented argument may be used to deduce later lines).

3.6.2   Conjunction

This is an easy section, but not so easy that one should completely ignore what is being said.

If our goal is to prove $A \wedge B$ (under local assumptions), break it into two subgoals: prove A, then prove B (or vice versa).

In all proof outlines, statements that are assumed or proved or proved under local assumptions have line numbers. The line numbers in the sample outlines are rather arbitrary, suggesting places where many missing lines would be present in an actual proof.

Goal: $A \wedge B$

Part I, Goal A:

proof steps ...

(25): A

Part II, Goal B:

proof steps: ...

(37): B

(38): $A \wedge B$ (lines 25, 37)

This is an example of a rule that can be applied at any time (in the absence of formal goal comments),

(25): A

proof steps: ...

(37): B

proof steps ...

(108): $A \wedge B$ (lines 25, 37)

If we have a local conclusion $A \wedge B$, we can draw further local conclusion A and B.

(2): $A \wedge B$

proof steps: ...

(102): A (line 2)

or

(2): $A \wedge B$

proof steps: ...

(102): B (line 2)

are good proof structures.

So, either as a goal or a local conclusion we can just break a conjunction into its two parts.

This can be phrased in a different way in terms of arguments.

$$\frac{A \quad B}{A \wedge B}$$

is a valid argument (this is how to prove a conjunction). This exemplifies a notation for arguments: write the premises above a horizontal bar and the conclusions below it. A linear notation is

$$A, B \vdash A \wedge B.$$

The rules for using a conjunction are just as simple:

$$\frac{A \land B}{A}$$

and

$$\frac{A \land B}{B}$$

are valid arguments.

3.6.3   Implication

If our goal is to prove $A \to B$, our strategy is to introduce a new local assumption A, and adopt B as a new goal. As soon as we prove B, we abandon the assumption A (and any local conclusions that we used A to prove) and can add $A \to B$ as a new local conclusion.

Goal: $A \to B$

> **Assume (14):** A
> **Goal:** B
> **proof steps:** …
> **(31):** B

(32): $A \to B$ deduction 14-31 (I used to say that closing lines were optional; I now always write them and require them; and I used to suggest restarting the numbering, which I no longer advise. Notice that this is called the rule of deduction).

If we have a local conclusion $A \to B$ and we also have a local conclusion A, we can draw the further local conclusion B. This is the rule of *modus ponens*.

(19): $A \to B$

proof steps: ...

(87): A

proof steps: ...

(256): B (modus ponens, lines 19,87)

> I'm not worried by order here

(19): A

proof steps: ...

(87): $A \to B$

proof steps: ...

(256): B (modus ponens, lines 19,87)

> is also fine.

If we just have the local conclusion $A \to B$, this can motivate the introduction of a new goal "prove A", the advantage of this goal being that once we meet it we can also conclude B. This amounts to adding a comment to one of the proof fragments above.

(19): $A \to B$

proof steps: ...

Goal: A (for the sake of getting B)

proof steps: ...

(87): A

(88): B (modus ponens, lines 19,87)

The fact that we prove B right after A in this sample is due to the indicated motivation: we are proving A in order to get B. There is no indentation at the goal because no assumption is being introduced.

These are the basic "forward" strategies for implication. We will introduce other strategies derived from the equivalence of an implication $A \to B$ with its contrapositive $\neg B \to \neg A$ in a later subsection.

These ideas can be phrased in terms of arguments. The usual strategy for using an implication is easy to phrase:

$$\frac{\begin{array}{c} A \\ A \to B \end{array}}{B}$$

is a valid argument: this expresses the rule of *modus ponens*.

The main strategy for proving an implication is a bit trickier to express. If

$$\frac{\begin{array}{c} A \\ H_1 \\ \ldots \\ H_n \end{array}}{B}$$

is a valid argument, then

$$\frac{\begin{array}{c} H_1 \\ \ldots \\ H_n \end{array}}{A \to B}$$

is a valid argument, where the H_i's are other background assumptions you may be using. This expresses the same idea given above.

The alternative approach suggested for using an implication as a hypothesis by itself is expressible in terms of arguments as well.

To prove that

$$\frac{\begin{array}{c} A \to B \\ H_1 \\ \ldots \\ H_n \end{array}}{C}$$

is a valid argument, prove that

$$H_1$$
$$\dots$$
$$H_n$$
$$\frac{}{A}$$

and

$$B$$
$$H_1$$
$$\dots$$
$$H_n$$
$$\frac{}{C}$$

are valid arguments.

3.6.4 Biconditional

To prove a biconditional $A \leftrightarrow B$ (under local assumptions) falls into two parts: first, prove $A \rightarrow B$ (assume A and prove the subgoal B); second, prove $B \rightarrow A$ (assume B and prove the subgoal A; the earlier assumption A was of course abandoned at the end of the first part).

Goal: $A \leftrightarrow B$

Part I: Assume(25): A

    Goal: B

    proof steps $\dots$

    (92): B

Part II Assume (93): B (I used to suggest restarting the numbering here because the previous block won't be used again, but the closing line will reference that block.)

    Goal: A

    proof steps $\dots$

    (134): A

(135): $A \leftrightarrow B$ biconditional introduction, 25-92, 93-134.

If we have a local conclusion $A \leftrightarrow B$, we can further conclude $A \rightarrow B$ and $B \rightarrow A$. This means for example that if we have $A \leftrightarrow B$ and A we can conclude B, and also that if we have $A \leftrightarrow B$ and B we can conclude A.

(19): $A \leftrightarrow B$

proof steps: ...

(87): A

proof steps: ...

(256): B (modus ponens for biconditional, lines 19,87)

    or

(19): $A \leftrightarrow B$

proof steps: ...

(87): B

proof steps: ...

(256): A (modus ponens for biconditional, lines 19,87)

The biconditional and the other line used in these rules might appear in the other order.

More rules for use of the biconditional will be given below.

3.6.5   Homework Assigned 9/9/2019, due 9/18/2019

This is assigned Monday 9/9/2019 and due Wednesday 9/18/2019.

It is worth noting that I have made some changes in my nomenclature for rules and my line numbering conventions since 2016; I am probably going to edit these notes to reflect such changes and will tell you when a new version has been posted. The 2018 version of my manual of logical style is more uptodate in this respect: there is a link to it on your class page.

1. Notice that the operators $\wedge$ and $\vee$ are associative, in the sense that $(A \wedge B) \wedge C$ is logically equivalent to $A \wedge (B \wedge C)$ and $(A \vee B) \vee C$ is logically equivalent to $A \vee (B \vee C)$.

   Prove using a truth table that $P \rightarrow (Q \rightarrow R)$ is not logically equivalent to $(P \rightarrow Q) \rightarrow R$. State a specific assignment of truth values for P, Q, R under which these two statements have different truth values.

   Prove using a truth table that $P \leftrightarrow (Q \leftrightarrow R)$ **is** logically equivalent to $(P \leftrightarrow Q) \leftrightarrow R$.

   Since $\leftrightarrow$ is associative, we can write $P \leftrightarrow Q \leftrightarrow R$ without parentheses. Briefly state under what conditions this statement is true (they might be surprising, but you can read them off the truth table).

   Extra credit: Give a brief explanation of the conditions under which $P_1 \leftrightarrow P_2 \leftrightarrow \ldots \leftrightarrow P_n$ is true for different values of n. The meaning is **not** "all the P_i's are equivalent".

2. Write an expression for something for which a truth table is given. Then do the same thing for implication:

   Write an expression using just $\neg$, $\vee$, $\wedge$ and letters for the operation represented by this truth table:

P	Q	R	???
T	T	T	F
T	T	F	T
T	F	T	F
T	F	F	F
F	T	T	T
F	T	F	F
F	F	T	T
F	F	F	F

28

Write an expression using just $\neg$, $\vee$, $\wedge$ and letters equivalent to $P \to Q$. Write the one you would actually read from the truth table: there is of course a shorter one.

Notice that the assignment continues on the next page!

3. Write a proof in the style developed in class of the biconditional statement
$$(A \to (B \wedge C)) \leftrightarrow ((A \to B) \wedge (A \to C))$$

This is **obviously** valid: this is an exercise in carefully writing out all the steps! Do you think you can prove
$$((A \wedge B) \to C) \leftrightarrow ((A \to C) \wedge (B \to C))?$$

4. We discussed in class how to show that the neither. . .nor. . . operation with the truth table

P	Q	$P|Q$
T	T	F
T	F	F
F	T	F
F	F	T

can be used to express anything represented by a truth table, by showing how to express $\neg P$, $P \wedge Q$ and $P \vee Q$ in terms of $|$ and letters alone.

Do the same for the NAND operation with the truth table

P	Q	$P \downarrow Q$
T	T	F
T	F	T
F	T	T
F	F	T

Show how to express implication $P \to Q$ and the biconditional $P \leftrightarrow Q$ in terms of $|$ and letters alone and in terms of $\downarrow$ and letters alone.

5. This exercise is related to the verification of the rule of constructive dilemma which I did in class, which is discussed in detail in the 2018 version of the manual of logical style (there is a link to the manual of logical style on the class page: the discussion is on pp. 28-30 of the manual).

Verify the rule of *destructive dilemma*

$$P \to R$$
$$Q \to S$$
$$\frac{\neg R \vee \neg S}{\neg P \vee \neg Q}$$

I give you the first few lines

(1): $P \to R$

(2): $Q \to S$

(3): $\neg R \vee \neg S$

Goal: $\neg P \vee \neg Q$

Hints: you could prove this either by cases or by alternative elimination. If you prove it by alternative elimination, do notice that the hypothesis will be $\neg\neg P$ not $\neg P$ (and you can get P from this right away). No matter which way you prove it, you will definitely want to use the rule of modus tollens one or more times. This is very similar to the constructive dilemma proof in either case, but just a little more indirect.

3.6.6 Disjunction

The strategies given for disjunction have limitations: more practical disjunction strategies using negation are given below.

The first rule we give for using a disjunction which appears as a local conclusion is "proof by cases". This strategy is very important to understand: it is constantly used in actual proofs written in English.

proof by cases: If you are given a local conclusion $A \vee B$ and aim to prove a goal C, the proof has two parts.

In Part I, assume A and adopt C as your goal.

In Part II, assume B and adopt C as your goal.

Here is a proof outline:

(42): $A \vee B$

proof steps ...

Goal: C

> **Case I (from (42)), Assume (53):** A (I'm tempted to call these assumptions 42a and 42b)
>
> **Goal:** C
>
> **proof steps** ...
>
> **(97):** C
>
> **Case II (from (42)), Assume (98):** B (anything proved in Case I has been discarded)
>
> **Goal:** C
>
> **proof steps** ...
>
> **(186):** C

(187): C (proof by cases)

There are a number of different strategies for proving a disjunction $A \vee B$. The ones which do not use negation appear rather limited.

1. To prove $A \vee B$, prove A. If you can do this, it works.

   Goal: $A \vee B$

   Goal (more specific): A

   proof steps: ...

   (43): A

   (44): $A \vee B$

2. To prove $A \vee B$, prove B. If you can do this, it works.

   Goal: $A \vee B$

   Goal (more specific): B

   proof steps: ...

   (43): B

   (44): $A \vee B$

These strategies here can be restated as "forward" rules: if you have the local conclusion A, you can further conclude $A \vee B$ (for any B: which B you choose will probably be guided by what you are trying to prove). If you have the local conclusion B, you can further conclude $A \vee B$ (for any A: which A you choose will probably be guided by what you are trying to prove).

These are both called rules of addition.

(32): A

proof steps . . .

(89): $A \lor B$ (addition, line 32) [B here can be anything: presumably we look at our current goals to decide what we need].

   or

(32): B

proof steps . . .

(89): $A \lor B$ (addition, line 32) [A here can be anything: presumably we look at our current goals to decide what we need].

You should not expect to have to prove a disjunction by deciding which of the alternatives is true. $A \lor \neg A$ is an easy example of a logical truth which is clearly valid and clearly cannot be decided by proving one of the disjuncts.

3.6.7 The First Example

Prove:

$$((P \land Q) \to R) \leftrightarrow (P \to (Q \to R))$$

The statement is a biconditional so the proof has two parts, one to prove $((P \land Q) \to R) \to (P \to (Q \to R))$ and one to prove $(P \to (Q \to R)) \to ((P \land Q) \to R)$.

Part I: Goal: $((P \land Q) \to R) \to (P \to (Q \to R))$

We apply the strategy for proving an implication.

Assume (1): $((P \land Q) \to R)$
Goal: $(P \to (Q \to R))$

We apply the strategy for an implication again.

Assume (2): P
Goal: $Q \to R$

Assume (3): Q
Goal: R

The goal now has no structure left: we have to do something different. Assumption 1 is an implication with hypothesis $P \land Q$.
We use the word Lemma for a partial result we prove as a means to our final result.

Lemma (4): $P \land Q$: because we have P (assumption 2) and Q (assumption 3).

Conclusion (5): Now the goal R follows, by the rule of modus ponens using assumption (1) and conclusion (4). This completes the proof of the goal of Part I, though I do supply some closing lines.

6: $Q \to R$ deduction, lines 3-5

7: $((P \land Q) \to R)$ deduction, lines 2-6

Part II: Goal: $(P \rightarrow (Q \rightarrow R)) \rightarrow ((P \wedge Q) \rightarrow R)$

We are proving an implication. All assumptions from Part I are now abandoned, so we can restart the numbering (though we differentiated the labels with b for the sake of the line references in our closing line).

Assume (1b): $(P \rightarrow (Q \rightarrow R))$

Goal: $((P \wedge Q) \rightarrow R)$

Assume (2b): $P \wedge Q$

Goal: R

Lemmas: We can immediately break assumption (2b) into lemma (3b) P and lemma (4b) Q.

Lemma (5b): We can conclude $Q \rightarrow R$ using modus ponens with assumption (1b) and lemma (3b).

Conclusion (6b): We can conclude R, our goal, using modus ponens with lemma (5b) and lemmma (4b). This completes the proof of Part II and of the whole theorem: we do supply closing lines tidying this up.

It is worth noting that we could replace the simple propositional letters P, Q, R here with letters A, B, C and the proof would go in exactly the same way and prove something more general, because A, B, C might stand for more complicated sentences of our logic.

7b: $((P \wedge Q) \rightarrow R)$ deduction, 2b-6b.

the Main Goal proved: $((P \wedge Q) \rightarrow R) \rightarrow (P \rightarrow (Q \rightarrow R))$ biconditional introduction, 1-7, 1b-7b. We could have proved the two implications and referenced them instead of the blocks.

3.7 Strategies using Negation

3.7.1 Negation

If we have a goal $\neg A$, the strategy for proving it (under local assumptions) is to assume A and adopt the new goal $\bot$ (contradiction) (once we establish this goal we immediately abandon the assumption A and any conclusions we drew using it with horror!)

The line numbers in this or any example are arbitrary.

Goal: $\neg A$

    Assume (33): A (for the sake of a contradiction)

    Goal: $\bot$ (we try to prove something ridiculous from A)

    proof steps Some proof steps. . .

    (57): $\bot$

(58): $\neg A$ negation introduction 33-57 (notice that this rule is called *negation introduction*).

If we have a local conclusion $\neg A$ and we also have the local conclusion A, we can draw the conclusion $\bot$ (we hope that this will only happen under nontrivial local assumptions!)

Notice that if we have A as a local assumption and wish to prove $\neg\neg A$, our strategy tells us to assume $\neg A$ and adopt $\bot$ as our new goal, which we can immediately meet because we have assumed both A and $\neg A$.

If we just have a local conclusion $\neg A$, a strategy which is sometimes useful is to adopt the goal A: if you can prove A, you get the further conclusion $\bot$, from which anything follows.

If we have a local conclusion $\neg\neg A$, we can draw the further conclusion A. We call this the rule of double negation.

This motivates the strategy "proof by contradiction", which is the only logical strategy which can be applied to a statement of any form at all. Suppose we want to prove A. We assume $\neg A$ for the sake of argument and adopt contradiction as our goal. If we succeed, we have proved $\neg\neg A$ by the strategy for proving negative statements, and we can draw the further conclusion A by the rule of double negation. Notice that this is not quite the same as our strategy for proving negative statements, though it uses it.

The justification for proof by contradiction:

Goal: A

Rewritten Goal: $\neg\neg A$ (A is equivalent to its double negation; we will not write this line in a proof by contradiction)

> **Assume(27):** $\neg A$ (we apply the standard negation strategy to the rewritten goal)
>
> **Goal:** $\bot$
>
> **some proof steps** some proof steps. . .
>
> **(102):** $\bot$

(103): $\neg\neg A$ negation introduction 27-102 [this line will not appear in an actual proof by contradiction]

(104): A (double negation, line 103)

We can drop lines involving double negations from the proof above to get the schematic outline of a proof by contradiction. The version above shows that proof by contradiction is justified by our other strategies (it is not really a new strategy).

Goal: A

> **Assume(27):** $\neg A$ (we assume that A is not true for the sake of a contradiction)
>
> **Goal:** $\bot$
>
> **some proof steps** some proof steps. . .
>
> **(102):** $\bot$

(103): A (proof by contradiction 27-102 (this rule can also be called *reductio ad absurdum*).

3.7.2 The Absurd

To prove $\bot$ (a contradiction), prove A and $\neg A$ for some statement A. We hope that we can only do this under local assumptions.

A sample bit of proof:

(32): A

proof steps ...

(71): $\neg A$

(72): $\bot$ (contradiction, lines 32,71)

If you have the local conclusion $\bot$, you may further conclude any statement A that you like (a false statement implies anything).

(112): $\bot$

(113): B (where B is anything at all, presumably our current goal).

3.7.3　Contrapositive

We first prove a theorem then introduce some new rules.

Theorem: $(P \to Q) \leftrightarrow (\neg Q \to \neg P)$

The theorem is a biconditional, so its proof can be expected to have two parts.

Part I: Goal: $(P \to Q) \to (\neg Q \to \neg P)$

Assume (1): $P \to Q$

Goal: $\neg Q \to \neg P$

Assume (2) $\neg Q$

Goal: $\neg P$

Now we use the proof strategy for a negative goal.

Assume (3): P

Goal: $\bot$ (contradiction)

Lemma (4): Q, by modus ponens using assumption (1) and assumption (3)

(5): $\bot$, because we have lemma (4), Q and assumption (2), $\neg Q$. This completes the proof of Part I, once we write closing lines.

(6): $\neg P$ neg intro 3-5

(7): $\neg Q \to \neg P$ deduction 2-6

Part II: Goal: $(\neg Q \to \neg P) \to (P \to Q)$

Assume (1b): $(\neg Q \to \neg P)$

Goal: $P \to Q$

Assume (2b): P

Goal: Q

I am going to prove Q by contradiction. Its the last resort: there doesn't appear to be anything else to do.

Assume (3b): $\neg Q$

Goal: $\bot$

Lemma (4b): $\neg P$: by modus ponens using assumption (1b) and assumption (3b).

(5b): ⊥ (contradiction), by assumption (2b) P and lemma (4b) $\neg P$. This completes the proofs of all goals back to Part II, and of the entire theorem – once we write the closing lines which follow. Note that we could replace the simple propositions P, Q, R with A, B, C and the proof would go in the same way and be more general (and indeed we need this generality below).

(6b): Q reductio ad absurdum, 3b-5b

(7): $P \to Q$ deduction 2b-6b

(8): the main theorem by biconditional introduction, 1-7, 1b-7b

The equivalence of an implication and its contrapositive motivates a different strategy for proving an implication and a different way of using an implication.

If we want to prove $A \to B$, we can make the assumption $\neg B$ and aim for the subgoal $\neg A$. [The reason that this works is that we can then conclude $\neg B \to \neg A$ by the direct proof strategy for implication, then use the theorem above to further conclude $A \to B$, but you do not need to write this out when you use this strategy].

A proof by contrapositive with the hidden justification in terms of the usual strategies:

Goal: $A \to B$

Rewritten Goal: $\neg B \to \neg A$

Assume(13): $\neg B$ (justified by our usual strategy for proving an implication, applied to the rewritten goal)

Goal: $\neg A$

proof steps: . . .

(25): $\neg A$

(26): $\neg B \to \neg A$ ded 13-25

(27): $(A \to B) \leftrightarrow (\neg B \to \neg A)$ [proved above]

(28): $A \to B$ (lines 26,27, one of the flavors of modus ponens for a biconditional)

This justifies the following derived strategy, which is what we will really write when we want to prove an implication indirectly:

Goal: $A \to B$

    Assume(13): $\neg B$

    Goal: $\neg A$

    proof steps: ...

    (25): $\neg A$

(26): $A \to B$ (indirect proof of an implication): indirect proof 13-25

If we have drawn the local conclusion $A \to B$, and also the local conclusion $\neg B$, we can further conclude $\neg A$ (this is the rule of *modus tollens*.

Here is the justification for *modus tollens* with the hidden extra steps to make it work with our basic strategies. You will not write these extra steps when using *modus tollens* as a rule!

(34): $A \to B$

proof steps ...

(82): $\neg B$

proof steps: ...

(117): $(A \to B) \leftrightarrow (\neg B \to \neg A)$ [proved above]

(118): $\neg B \to \neg A$ (m.p., lines 34, 117)

(119): $\neg A$ (m.p., lines 82, 118)

This justifies the more compact use of modus tollens as an independent rule:

(34): $A \to B$

proof steps ...

(82): $\neg B$

proof steps: ...

(117): $\neg A$ (m.t., lines 34, 82)

41

3.7.4 Rules with Disjunction and Negation

As we commented above, you should not expect to have to prove a disjunction by deciding which of the alternatives is true. We will see a very simple example soon where this is clearly impossible. A truth table can be used to show that $A \lor B$ is logically equivalent to $\neg A \to B$, and also to $\neg B \to A$: application of the strategies for implication to these equivalent forms motivate more general rules for disjunction which will involve negations. We give two different forms of the rules, which swap the roles of the variables.

  We discuss in this section how the rules given here for disjunction can be used to justify proof by cases and addition. It is also true that the rules of proof by cases and addition together with the rules of negation can be used to derive the rules in this section; the appeal to truth tables is purely a matter of convenience.

  1. This strategy depends on the equivalence of $A \lor B$ with $\neg A \to B$ (which you can verify using a truth table).

     To prove $A \lor B$, assume $\neg B$, and adopt the subgoal A.

     Goal: $A \lor B$

       Assume(32): $\neg B$
       Goal: A
       proof steps: ...
       (87): A

     (88): $A \lor B$ alternative elimination 32-87

2. This strategy depends on the equivalence of $A \lor B$ with $\neg B \to A$ (which you can verify using a truth table).

To prove $A \lor B$, assume $\neg A$, and adopt the subgoal B.

Goal: $A \lor B$

    Assume(32): $\neg A$
    Goal: B
    proof steps: ...
    (87): B

(88): $A \lor B$ alternative elimination, 32-87

Notice that these rules cover the rules without negation: for example, if we can prove $A \lor B$ by proving A, we can also prove $A \lor B$ by assuming $\neg B$ and then proving A in the same way without using the assumption. It is also true that the rules for disjunction with negation can be derived from the rules for disjunction without negation (with the help of proof by contradiction). This is in the homework.

Something I have seen in earlier classes is perfectly correct but redundant proofs of a disjunction with two parts, one using each of the last two techniques. This is not necessary: if you have proved a disjunction by one of these methods, it is proved.

An easy theorem:

Theorem (Excluded Middle): $A \lor \neg A$

Notice that you can't possibly prove this using the strategies for proving a disjunction without negation, as neither A nor $\neg A$ (simple statements about which we have no information) are provable.

Assume (1): $\neg A$ (assume the negation of the first disjunct).

Goal: $\neg A$ (the goal is the second disjunct).

Conclusion: There is nothing to do, as the goal is the assumption – we are done!

(2): $A \lor \neg A$ alternative elimination 1-1.

This justifies a rule: at any point in a proof you may introduce a line of the form $A \vee \neg A$ with the justification "excluded middle".

We give two rules for using a disjunction $A \vee B$ which we have as a local conclusion, based on the equivalences with $\neg A \rightarrow B$ and $\neg B \rightarrow A$. A general name for these rules is *disjunctive syllogism*.

disjunctive syllogism (1): Given $A \vee B$ and $\neg A$, further conclude B. (This uses the equivalence of $A \vee B$ and $\neg A \rightarrow B$ and *modus ponens*).

> **(31):** $A \vee B$
>
> **proof steps** ...
>
> **(45):** $\neg A$
>
> **proof steps** ...
>
> **(117):** B (disjunctive syllogism, lines 31, 45).

The disjunction and the negation could be in the other order.

disjunctive syllogism (2): Given $A \vee B$ and $\neg B$, further conclude A. (This uses the equivalence of $A \vee B$ and $\neg B \rightarrow A$ and *modus ponens*).

> **(31):** $A \vee B$
>
> **proof steps** ...
>
> **(45):** $\neg B$
>
> **proof steps** ...
>
> **(117):** A (disjunctive syllogism, lines 31, 45).

The disjunction and the negation could be in the other order.

Two further rules which combine disjunctive syllogism and double negation can also be called special cases of disjunctive syllogism.

disjunctive syllogism (3): Given $\neg A \vee B$ and A, further conclude B. (replace A with $\neg \neg A$ and use the first disjunctive syllogism rule).

disjunctive syllogism (4): Given $A \vee \neg B$ and B, further conclude A. (replace B with $\neg \neg B$ and use the second disjunctive syllogism rule).

The example theorem I prove for this section shows why the rule of Proof by Cases can be justified using only disjunctive syllogism. It is also a nice demonstration of the contrapositive proof methods.

Theorem: $((A \to C) \land (B \to C)) \to ((A \lor B) \to C)$

> **Assume (1):** $(A \to C) \land (B \to C)$
> **Goal:** $(A \lor B) \to C$
>> **Assume (2):** $A \lor B$
>> **Goal:** C
>> I will prove this by contradiction.
>> **Assume (3):** $\neg C$
>> **Goal:** $\bot$ (contradiction).
>> **Lemmas:** From assumption (1) we get lemma (4) $A \to C$ and lemma (5) $B \to C$
>> **Lemma (6):** $\neg A$, from (3), (4) by modus tollens.
>> **Lemma (7):** $\neg B$, from (3), (5) by modus tollens.
>> **Lemma (8):** B, from (2), (6) by disjunctive syllogism.
>> **(9):** $\bot$, from (7) and (8). This completes all goals and the proof, after closing lines...
> **10:** C reduction ad absurdum 3-8

Notice that Part I of a proof by cases of C from $A \lor B$ is a proof of $A \to C$, and Part II of a proof by cases of C from $A \lor B$ is a proof of $B \to C$: this theorem tells us that if we have $A \lor B$ and we have these two proofs, we also have $(A \lor B) \to C$ and then C by *modus ponens*.

3.8 Propositional Logic in Marcel

[NOTE: I'm going to write some fresh remarks here]

3.9 Propositional Logic Style Sheet

This is the style sheet which was attached to Test I the first time I taught from these notes. You may expect something similar.

Not all of these are relevant to the assigned proofs; you need to recognize which rules and strategies are appropriate.

Conjunction (and): To prove $A \wedge B$, prove A (part 1), then prove B (part 2). The parts are not cases, and you will lose credit if you call something a proof by cases which isn't one.

From $A \wedge B$, deduce A. From $A \wedge B$, deduce B. You need to explicitly break apart assumptions or lemmas which are "and" statements to use their components; you will lose credit if you don't.

Disjunction (or): To prove $A \vee B$, assume $\neg A$ and adopt the new goal B [or assume $\neg B$ and adopt the new goal A; you do not need to do both].

From A, deduce $A \vee B$. From B, deduce $A \vee B$ (rule of addition or (as I have called it) weakening).

From $A \vee B$ and $\neg A$, deduce B. From $A \vee B$ and $\neg B$, deduce A. This is disjunctive syllogism.

To deduce a conclusion C from an assumption or lemma $A \vee B$, use *proof by cases*: in the first part (case 1) assume A and prove C; in the second part (case 2) assume B and prove C.

Implication (if): To prove $A \to B$, assume A and adopt the new goal B.

alternative strategy: to prove $A \to B$, assume $\neg B$ and adopt the new goal $\neg A$.

Given A and $A \to B$, deduce B (*modus ponens*).

Given $\neg B$ and $A \to B$, deduce $\neg A$ (*modus tollens*).

If you have an assumption $A \to B$ you may want to try proving A (so that you can further conclude B).

Negation (not): To prove $\neg A$, assume A and try to prove $\bot$ (contradiction).

From $\neg\neg A$ deduce A (double negation).

To prove any statement A, assume $\neg A$ and try to prove $\bot$. This is proof by contradiction.

From A and $\neg A$, deduce $\bot$.

If you have a negative assumption $\neg A$, the commonest way to use it is to wait until your goal is a contradiction, then try to prove A to get the contradiction.

From $\bot$ you may deduce anything.

4 Logic of Quantifiers

In this section we increase the power of our logic by introducing translations of a different kind of logical construction, which is signalled in English by words like "all", "some", "there exists" and some other phrases.

4.1 Atomic sentences analyzed

So far we have considered simple sentences as mere letters. However, it must be noted that they actually have more structure. A mathematical sentence like $x < y$ has a structure very like that of an English sentence such as "John loves Mary". In these sentences we have (in terms of English grammar) a subject, object and verb. In logic, we call subject and object "arguments" (a confusing word we won't use often): they are the things the sentence is talking about. The verb we call the "predicate".

We can also have sentences with one argument, such as "x is prime" or "the Sun is green" (no reason our sentences have to *true*). Three arguments may occur occasionally, as in "John handed the money to Sarah" or "P is between Q and R".

Sentences may contain complex names, as in $x + y = z$ or "The owner of the car paid John". This will cause no particular logical problems for us.

4.2 Quantifiers

The logical constructions we now introduce are called *quantifiers*.

"For all x, $P(x)$", in symbols $(\forall x.P(x))$ says that $P(a)$ is true for any object a. This is the *universal quantifier*. It can also be written in English, "for any x, $P(x)$", "for each x, P(x)", and universal quantifiers can appear indirectly in natural or mathematical English in various ways.

"For some x, $P(x)$", in symbols $(\exists x.P(x))$ says that $P(a)$ is true for some object a. This can also be expressed as "There exists x such that $P(x)$". This is the *existential quantifier*.

Quantifiers are very seldom really over absolutely everything. The domain (for example, real numbers) may be understood from context but can also be expressed explicitly.

$(\forall x \in A.P(x))$ means $(\forall x.x \in A \rightarrow P(x))$. This can be read "for all x in the set A, $P(x)$".

$(\exists x \in A.P(x))$ means $(\exists x.x \in A \wedge P(x))$. Notice that "and" is used instead of "implies". This can be read "for some x in the set A, $P(x)$", or "there is an x in A such that $P(a)$".

These constructions are called "restricted quantifiers".

An example of the importance of understanding what the context is: $(\exists x.x + x = 1)$ is true if the domain of discourse is the real numbers, and false if it is the integers or natural numbers.

Quantified sentences may have deceptive surface grammatical forms in English. The classical philosophical example "All men are mortal" is read $(\forall x.x$ is a man$\rightarrow x$ is mortal). The surface form of the sentences in English does signal the universal quantifier, but it takes practice to see the implication.

Multiple quantifiers require care. We exhibit two sentences in English:

Everyone loves someone

Someone is loved by everyone

These do not seem to mean the same thing, and this is odd, since "John loves Mary" and "Mary is loved by John" certainly mean the same thing.

The logical analysis:

$(\forall x.(\exists y.x \, L \, y))$ for "Everyone loves someone"

$(\exists y.(\forall x.y$ is loved by $x))$, that is, $(\exists y.(\forall x.x \, L \, y))$ for "Someone is loved by everyone".

The difference is just the order of the quantifiers.

The first sentence is true and the second sentence is false in a mini-world where Alice and Bob love each other and Sue and Doug love each other. $(\forall x.(\exists y.x \, L \, y))$ is true because for any specific person a, $(\exists y.a \, L \, y)$ is true. The exact person that a loves depends on who a is. The second sentence

$(\exists y.(\forall x.x \, L \, y))$ is false in this mini-world because there is no choice of a which makes $(\forall x.x \, L \, a)$ true: everyone loves someone, but not the same person. In the first sentence, y can be chosen in a way which depends on x, while in the second a single choice for y must be made once and for all.

These issues arise in the calculus example we did on day 1, which can be written in our notation

$$(\forall \epsilon > 0.(\exists \delta > 0.(\forall x.0 < |x - a| < \delta \rightarrow |f(x) - L| < \epsilon))).$$

The order of the quantifiers signals that the choice of δ can depend on the choice of ϵ.

A mathematical example of the order of quantifiers making a difference: $(\forall x.(\exists y.x < y))$ says "for every number x there is a larger number y" (true) while $(\exists y.(\forall x.x < y))$ says that there is a fixed number y greater than all numbers x (a largest number), which is false.

It can be a useful exercise to translate basic things that we usually say in mathematical English completely into mathematical notation, now that we have a more powerful logical vocabulary. For example, "x is prime" can be expressed as

$$x \in \mathbb{N} \wedge p > 1 \wedge (\forall a.a|p \rightarrow a = p \vee a = 1).$$

Things can be taken all the way down to basic operations if we further define $x|y$ as $(\exists a \in \mathbb{N}.ax = y)$.

Something to be aware of: in a sentence $(\forall x.P(x))$ or $(\exists x.P(x))$ does not refer to any specific object x. This is generally clear where a universal sentence is concerned but less clear for an existential statement. $(\exists x \in \mathbb{R}.x + x = 1)$ is true, and moreover, there is a unique value $x = \frac{1}{2}$ which makes it true. But a statement of this form does not necessarily talk about a specific object: consider $(\exists x \in \mathbb{R}.x^2 = 1)$.

Negations of quantified sentences need to be handled carefully, just as we saw with negations of propositional sentences.

$\neg(\forall x.P(x))$, it isn't the casethat $P(x)$ for every x, is equivalent to $(\exists x.\neg P(x))$, not to $(\forall x.\neg P(x))$, which says that $P(x)$ is not true for any x at all. To show that a universal statement is false, you only need to find one counterexample.

$\neg(\exists x.P(x))$, there doesn't exist an x such that $P(x)$, is equivalent to $(\forall x.\neg P(x))$ (for any x, $P(x)$ is not true).

4.3 Proof Strategies

Just as with propositional logic constructions here, the point of our logical analysis is to guide us in constructing proofs.

4.3.1 Universal Quantifier

To prove a universal statement $(\forall x.P(x))$:

Goal: $(\forall x.P(x))$

> **Let** k be an arbitary object.
> **Goal:** $P(k)$

There are some variations of this.
This strategy handles a universal quantifier over a restricted domain.

Goal: $(\forall x \in A.P(x))$

> **Let** k be an arbitary object.
> **Assume:** $k \in A$
> **Goal:** $P(k)$

This strategy combines the universal quantifier and implication strategies: this is a very common kind of statement.

Goal: $(\forall x \in A.P(x) \rightarrow Q(x))$

> **Let** k be an arbitary object.
> **Assume:** $P(k)$
> **Goal:** $Q(k)$

If we have proved a local conclusion $(\forall x.P(x))$, and if a is any name of an object, we can also conclude $P(a)$. This is called universal generalization.

If we have proved a local conclusion $(\forall x \in A.P(x))$ and also $a \in A$, we can conclude $P(a)$.

If we have proved a local conclusion $(\forall x \in A.P(x) \rightarrow Q(x))$ and also $P(a)$, we can conclude $Q(a)$. This is a combination of universal generalization and modus ponens.

If we have proved a local conclusion $(\forall x \in A.P(x) \rightarrow Q(x))$ and also $\neg Q(a)$, we can conclude $\neg P(a)$. This is a combination of universal generalization and modus tollens.

4.3.2 Existential Quantifier

To prove a statement of the form $(\exists x.P(x))$:

Goal: $(\exists x.P(x))$

> **Find** a suitable c. This will usually take some thought.
> **Goal:** $P(c)$

The version for a restricted quantifier:
To prove a statement of the form $(\exists x \in A.P(x))$:

Goal: $(\exists x.P(x))$

> **Find** a suitable c. This will usually take some thought.
> **Goal 1:** $c \in A$ (this often doesnt require much work)
> **Goal 2:** $P(c)$

If $(\exists x.P(x))$ is a local conclusion, it allows us to do a special maneuver: we can introduce a new name w (not used anywhere before in the proof) and assume $P(w)$ [and that is all we can assume about w]. If $(\exists x \in A.P(x))$ is a local conclusion, we can do the same thing with the additional assumption $w \in A$. If you use this logical strategy more than once, be sure you use a different name each time. The name w is not generally consider permanent: it expires at the end of the proof (so we can use the same letter for a similar purpose in a later proof).

4.4 An Example

This is an example from Math 187.

Convention: the domain of discourse is the integers.

Definition: "n is even" means $(\exists k.n = 2k)$

Definition: "n is odd" means $(\exists k.n = 2k - 1)$

Theorem: The product of two odd numbers is odd.

Logical Analysis: This means "For all x, y, if x and y are odd then xy is odd". In practice we don't need to write this out in symbols, but an understanding of the logical structure guides our proofs. ($\forall xy.x$ is odd $\wedge y$ is odd $\rightarrow xy$ is odd) is a symbolic expression for this: notice the abbreviation for two universal quantifiers.

Let a, b be arbitrary integers

Assume (1): a is odd

Assume (2): b is odd

Goal: ab is odd.

Comment: You should recognize a combination of the universal quantifier and implication proof strategies. Here I avoid using the same letters x, y for the arbitrary objects I introduce that appear in the original theorem. One *can* use x, y but I think this is better style.

expand assumption 1 using definition: (3) ($\exists z.a = 2z - 1$)

expand assumption 2 using definition: (4) ($\exists w.b = 2w - 1$)

introduce witness to (3): We can choose a j such that (5) $a = 2j - 1$

introduce witness to (4): We can choose a k such that (6) $b = 2k - 1$

Comment: I am spelling things out in insane detail here to make a point: remember that in class we basically wrote (5) and (6) right after assumptions (1) and (2): I want to point out here that expanding definitions is an important logical move, and also that we are using the official strategy for using an existentially quantified assumption.

Expand Goal using definition: Goal: ($\exists u.ab = 2u - 1$)

Find c such that $ab = 2c - 1$

Scratch Work: $ab = (2j - 1)(2k - 1) = 4jk - 2j - 2k + 1 = (4jk - 2j - 2k + 2) - 1 = 2(2jk - j - k + 1) - 1$, so let $c = 2jk - j - k + 1$ and we are done!

Conclusion: $ab = 2c - 1$, when c is defined as $2jk - j - k + 1$, by the scratch work. This completes the proof of our goal and of the theorem.

4.5   Homework Exercises for 9/18/2019, due 9/25/2019

These are due at class on 9/25/2019.

1. Prove $((P \to Q) \land (Q \to R)) \to (P \to R)$ using the natural deduction rules (this should be straightforward).

2. Prove $(\neg P \lor Q) \leftrightarrow (P \to Q)$ using natural deduction. Notice that this is the equivalence which we used to justify the rules for disjunction involving negation; I will be impressed if you can prove it without using the rules that combine disjunction and negation (this would require some reasoning by contradiction), but you are welcome to prove it using the full rules.

3. Prove the deMorgan law $\neg(P \land Q) \leftrightarrow \neg P \lor \neg Q$ using the natural deduction rules (no equational calculations please). We might have done this one as a class exercise...then your writeup should be excellent!

4. Prove that the sum of two odd numbers is even, using the same definitions and style as in my in-class proof that the product of two odd numbers is odd (which appears in the notes on the previous pages).

4.6   Remarks on Quantifier Logic in Marcel

Just a couple of remarks to support the optional Marcel lab.

The commands needed to handle the Marcel quantifier lab are as before, with one addition. There is something to be said about the syntax of the Marcel language with quantifiers, but you should be able to read it, and the setup commands for the problems in the optional lab will be supplied to you; you will not need to write such expressions yourself in this lab.

When rules are applied to quantifier statements, Marcel will introduce new variables of the shape `x_1` (letter followed by underscore followed by numeral) for "arbitrary objects" (the kind you introduce when proving a universal statement or introducing a witness to an existential assumption), or new variables of the shape `y$1` (a letter followed by a dollar sign followed by a number) for objects that you need to choose (the objects to plug into a universal hypothesis; the object you choose to show that an existential conclusion is true).

There is a command for giving a "dollar sign variable" a value: for example to replace `x$3` with `z_1` you issue the command

```
SU('x$3 := z_1')
```

The resemblance to an assignment command in a computer language is not an accident. You do need to put the "assignment statement" in quotes, as Python is processing it as a string.

5   Formal Arithmetic, or Arithmetic Made Difficult

We are now going to spend some time doing formal proofs in the arithmetic of natural numbers starting with a very restricted set of assumptions. It is an important part of formal proof to understand exactly what one is allowed to use, and here the assumptions one is allowed to use will be spelled out precisely (and will be uncomfortably few).

The reason I do this is that I want us to work with quantifier proof in a controlled environment, but not completely in the abstract: it is *too* abstract.

5.1   Basic Concepts and Axioms of Formal Arithmetic

All of our objects are natural numbers. We denote the set of natural numbers by $\mathbb{N}$; we will actually not need to refer to it in these sections as the universe is inhabited only by natural numbers. We have the following basic concepts: 0 (a particular natural number) and the operations $S(x)$ (successor of x), addition and multiplication.

1. 0 is a natural number (in symbols, $0 \in \mathbb{N}$).

2. If x and y are natural numbers, so are $S(x)$, $x + y$, and $x \cdot y$. ($\forall xy \in \mathbb{N}.S(x) \in \mathbb{N} \wedge x + y \in \mathbb{N} \wedge x \cdot y \in \mathbb{N}$).

3. 0 is not a successor. ($\forall x.S(x) \neq 0$). Here we understand that $x \neq y$ abbreviates $\neg x = y$. Here and in the following axioms we write our quantifiers unrestricted: we could write ($\forall x \in \mathbb{N}.S(x) \neq 0$) instead, but in this context we are only talking about natural numbers, so we can leave the restriction on our quantifiers implicit.

4. Numbers with the same successor are the same. ($\forall xy.S(x) = S(y) \rightarrow x = y$).

5. Let P(x) be any sentence about a natural number variable x. We assert $P(0) \wedge (\forall y.P(y) \rightarrow P(S(y))) \rightarrow (\forall x.P(x))$. This is a symbolic presentation of the familiar principle of mathematical induction. From an extremely technical standpoint, this is an infinite collection of axioms, one for each sentence $P(x)$. If we are also willing to talk about sets of natural numbers, we can state it as a single axiom: ($\forall A \in \mathcal{P}(\mathbb{N}).0 \in A \wedge (\forall y \in \mathbb{N}.y \in A \rightarrow S(y) \in A) \rightarrow A = \mathbb{N}$). We will not use the set formulation now but we might use it later. $\mathcal{P}(\mathbb{N})$ is a notation for the collection of all sets of natural numbers.

6. ($\forall x.x + 0 = x$)

7. ($\forall xy.x + S(y) = S(x + y)$)

8. ($\forall x.x \cdot 0 = 0$)

9. ($\forall x.x \cdot S(y) = x \cdot y + x$) Here we assume the usual order of operations.

The original axiom system of Peano used 1 as the initial natural number instead of 0. This requires that some axioms be rewritten: I suggest this as an exercise.

Peano allowed himself to use some set theory as part of his logic: this allowed him to define addition and multiplication cleverly and prove their properties using very basic set theory and the first five axioms. This is out of fashion: we now use a formulation with no sets and with axioms in effect expressing recursive definitions of addition and multiplication.

If we understand successor, addition and multiplication on natural numbers, we certainly see that all these statements are true. We need to be careful to realize that we are *not* given other familiar properties of the natural numbers. We do not yet know that $0 + x = x$ or that $x + y = y + x$, for example. The fun thing is that these familiar properties *do* follow from this seemingly very impoverished set of assumptions.

It is important to note that axioms 1 and 2 are never explicitly used here, because we regard all quantifiers as implicitly restricted to natural numbers anyway. Nonetheless they are important assumptions (these are called "closure axioms").

5.2   About Equality and Substitution

The study of equality is part of logic. Here we note that $(\forall x. x = x)$ is a logical principle (that quantifier is completely unrestricted), called the reflexive property of equality.

Let $P(z)$ be any sentence containing a variable z: let $P(x)$ for example be the result of replacing all occurrences of z with x. $(\forall xy. x = y \to (P(x) \leftrightarrow P(y)))$ is a logical principle (or one principle for each sentence), a version of the *principle of substitution*. We can replace equals with equals in any sentence. The reason I phrased it in terms of $P(z)$ was that we might not want to replace every occurrence of x with y. For example, suppose $P(z)$ is defined as $y = z$. Then substitution tells us that $(\forall xy. x = y \to (y = x \leftrightarrow y = y))$. Now we can prove the symmetric property of equality:

Goal: $(\forall xy. x = y \to y = x)$

> **Let** a, b be arbitrary objects
> **Assume (1):** $a = b$
> **Goal:** $b = a$

Lemma(2): $(\forall xy.x = y \rightarrow (y = x \leftrightarrow y = y))$ (principle of substitution, from above)

Lemma (3): $a = b \rightarrow (b = a \leftrightarrow b = b)$ (from (2) with $x := a; y := b$)

Lemma (4): $b = a \leftrightarrow b = b$ m.p. (1) and (3)

Lemma (5): $b = b$ reflexivity of equality

Conclusion: $b = a$ by the biconditional (4) and (5) (using one of the m.p. like rules for using a biconditional) This was our goal and the proof is complete.

We will *not* do formal proofs of properties of equality in the abstract like this. fear not.

We do mention another form of substitution. Suppose $A(x)$ is a numerical expression (not a statement), and suppose $x = y$. Then we have $A(x) = A(x)$ by reflexivity of equality, and $A(x) = A(y)$ by substitution (this is an example where our use of $P(z)$ in the definition of substitution makes sense: $P(z)$ here is $A(x) = A(z)$ and we have $P(x)$, which is $A(x) = A(x)$, implying $P(y)$, which is $A(x) = A(y)$: if we just took $A(x) = A(x)$ and replaced all x's with y's, we would get $A(y) = A(y)$, which is not what we want (though it is true and does follow). That $x = y$ implies $A(x) = A(y)$ is the important idea of "doing the same thing to both sides of an equation".

Another important property of equality is transitivity: $x = y$ and $y = z$ implies $x = z$. This also follows from reflexivity and substitution, but we won't prove it here.

Application of transitivity motivates the notation $x = y = z$, meaning $x = y \wedge y = z$, which generalizes to $x_1 = x_2 = x_3 = \ldots x_{n-2} = x_{n-1} = x_n$ meaning $x_1 = x_2 \wedge x_2 = x_3 \wedge \ldots \wedge x_{n-2} = x_{n-1} \wedge x_{n-1} = x_n$. If we have a chain of equalities like this and we have justified each $x_i = x_{i+1}$ then we have justified $x_1 = x_n$.

5.2.1 Equality style manual

Here I am going to list examples of expected and allowed justifications of lines using properties of equality. I show snippets of proof under each rule.

Reflexivity of equality

117: $2+2 = 2+2$ ref $=$

Substitution

> **12:** $A = B$
>
> **more lines not shown** ...
>
> **53:** $P(A)$ (notice this is any statement about A)
>
> **more lines not shown** ...
>
> **117:** $P(B)$ substitution using line 12 into line 53

Symmetry

> **12** $A = B$
>
> **more lines not shown** ...
>
> **32** $B = A$ symm = line 12

Transitivity order of the premises doesnt matter, if 22 and 72 were interchanged this would still work

> **22** $A = B$
>
> **more lines not shown** ...
>
> **72** $B = C$
>
> **more lines not shown** ...
>
> **117** $A = C$ trans = 22,72

Symmetry and Transitivity order of the premises doesnt matter. These are both "things equal to the same thing are equal to each other".

> **3** $A = B$
>
> **more lines not shown** ...
>
> **21** $A = C$
>
> **more lines not shown** ...
>
> **104** $B = C$ symm trans = 3,21
>
> or

3 $B = A$

more lines not shown ...

21 $C = A$

more lines not shown ...

104 $B = C$ symm trans = 3,21

Chained transitivity (and possibly symmetry) Chains of equations of any length may be handled in a single line.

57 $A = B$

more lines not shown ...

61 $D = E$

more lines not shown ...

72 $C = B$

more lines not shown ...

82 $C = D$

more lines not shown ...

99 $A = D$, chain of equations, lines 5,7,8,6 (notice that I put the premises in the correct order for the chain in the justification)

Doing the same thing to both sides

18: $A = B$

more lines not shown ...

53 $F[A] = F[B]$ both sides, line 18

(where $F[A]$ is any complex expression in which replacing some A's with B's will give $F[B]$)

This saves a line from the following approach to proving the same thing

18: $A = B$

more lines not shown ...

52 $F[A] = F[A]$ ref $=$

53 $F[A] = F[B]$ substitution into line 52 using line 18

Something I think is too short

12 $S(S(a)) = S(b)$ who knows why?

72 $S(S(a)) = S(b + 0)$ axiom 6, $x := b$, subs

I think that is too short. The instance of axiom 6 that you use is never written down at all. I want you to write

12 $S(S(a)) = S(b)$ who knows why?

71 $b = b + 0$ axiom 6, $x := b$

72 $S(S(a)) = S(b)$ substitution using line 71 into line 12

I may use the first style, with the extra remark subs to signal the cheat, in board work, but I will use the second one in the notes and I expect you to use the second one in homework.

5.3 First proofs: 2+2=4; we get the successor by adding one.

We would like to write $x + 1$ instead of $S(x)$. But we do not know what 1 is yet.

Definition: We define 1 as $S(0)$, 2 as $S(S(0))$, 3 as $S(S(S(0)))$, 4 as $S(S(S(S(0))))$, and so forth.

Theorem: $2 + 2 = 4$

(1) $2 + 2 = 2 + 2$

(2) $2 + 2 = 2 + S(S(0)))$ definition of 2

(3) $2 + S(S(0)) = S(2 + S(0))$ axiom 7 $x := 2; y := S(0)$

(4) $2 + 2 = S(2 + S(0))$ transitivity of equality 2,3

(5) $2 + S(0) = S(2 + 0)$ axiom 7 $x := 2; y := 0$

(6) $2 + S(S(0)) = S(S(2 + 0))$ sub into 4 using 5

(7) $2 + 0 = 2$ axiom 6 $x := 2$

(8) $2 + S(S(0)) = S(S(2))$ sub into 6 using 7

(9) $2 + 2 = S(S(S(S(0))))$ two different uses of definition of 2 (in opposite directions)

(10) $2 + 2 = 4$ definition of 4. The result is proved.

Theorem: $(\forall x.x + 1 = S(x))$

Let a be an arbitrarily chosen natural number.

Goal: $a + 1 = S(a)$

(1) $a + 1 = a + 1$ reflexivity of equality

(2) $a + 1 = a + S(0)$ definition of 1

(3) $a + S(0) = S(a + 0)$ axiom 7, $x := a$

(4) $a + 1 = S(a + 0)$ transitivity of equality lines 2 and 3

(5) $a + 0 = a$ axiom 6, $x := a$

Conclusion $a + 1 = S(a)$ substitution into 4 using 5, and this completes the proof.

I am giving examples here of the style I expect. I want examples of axioms 6 through 9 or similar theorems stated for specific values used as here (this is an example of how universal statements are used) and then substitutions using these equations indicated. The reader (me) is supposed to be able to identify where the substitution is being made. This is very painstaking and we will not do it through the whole course. We are paying attention to our thought processes...

Once we have this theorem, we have the right to write $x + 1$ instead of $S(x)$ if we wish. I do suggest that this is not a good idea until we have proved certain basic arithmetic theorems. We do for example have $x + (y + 1) = (x + y) + 1$ (this is axiom 7 rewritten using the theorem above) but we do not yet have associativity of addition. Until we do, I suggest that the above notation for axiom 7 might mislead us.

5.4 Induction presented as a proof strategy

In this system of formal arithmetic, most of the power is included in the axiom of mathematical induction. This principle is already familiar to you, and my presentation of it as a proof strategy similar to the proof strategies we presented above in propositional and quantifier logic should not really be surprising.

Goal: $(\forall x \in \mathbb{N}.P(x))$

> **Basis:** Goal: $P(0)$
>
> **Induction Step:** (Goal: $(\forall y.P(y) \rightarrow P(S(y)))$) writing this goal is optional.
>
> > **Let** k be an arbitary natural number.
> > **Assume (inductive hypothesis):** $P(k)$
> > **Induction Goal:** $P(S(k))$ [or $P(k+1)$]

This is exactly the strategy for proving $P(0) \wedge (\forall k.P(k) \rightarrow P(S(k)))$, and axiom 5 tells us that this implies $(\forall x.P(x))$. This statement is NOT part of the proof outline: you do not need to repeat it after every induction proof. It's just a remark about why the proof outline works.

We state the quantifiers explicitly in the proof outline because we will also use this principle in the more general setting of real analysis where not all objects are natural numbers.

5.5 Our first two induction proofs

Theorem: $(\forall x.S(x) \neq x)$

> **Basis:** Goal: $S(0) \neq 0$
>
> > **Basis checks out:** $S(0) \neq 0$ is axiom 3 with $x := 0$, and this is our goal.
>
> **Induction Step: Let** k be an arbitrarily chosen natural number.
> > **Assume (ind hyp):** $S(k) \neq k$
> > **Induction Goal:** $S(S(k)) \neq S(k)$
> > **(1)** $S(S(k)) = S(k) \rightarrow S(k) = k$ axiom 4, $x := S(k); y := k$

Conclusion: $S(S(k)) \neq S(k)$ by (1), (ind hyp) and modus tollens.

Comment: This is not a theorem we use terribly often. It is a proof involving induction on just one variable which is fairly easy to carry out and doesn't have weird features like the following one.

The propositional logic bit in this proof can be done in various ways.

Theorem: $(\forall x.x = 0 \vee (\exists y.S(y) = x))$

Comment: This theorem says that every number except 0 has a predecessor. By the way, axiom 4 tells us that a number has at most one predecessor; this theorem only says that if a number is not zero it has at least one predecessor.

Basis Step: Goal: $0 = 0 \vee (\exists y.S(y) = 0)$

Basis step checks out: $0 = 0$ is true by reflexivity of equality, and this is enough for this disjunction to be true. It is interesting to note that the other disjunct is always false by axiom 3, but this doesn't change the fact that the basis goal is proved.

Induction Step: Let k be an arbitrarily chosen natural number.

Assume (ind hyp): $k = 0 \vee (\exists y.S(y) = k)$

Induction Goal: $S(k) = 0 \vee (\exists y.S(y) = S(k))$

induction goal checks out: To show that this disjunction is true, it is sufficient to show that the second disjunct is true: $(\exists y.S(y) = S(k))$ ia true because $S(k) = S(k)$ is true: the appropriate value of y is k. It is interesting to note as we did in the basis step that the other disjunct is false by axiom 3, but this does not impair the proof, which is complete.

Comment: The second proof is very weird, and essentially unique among induction proofs, in make no use at all of the inductive hypothesis.

5.6 Commutativity of addition

Our aim is to prove $(\forall xy.x + y = y + x)$, the commutative law of addition.

We must be proving it by math induction, as we have no other way to do it!

We will prove this by induction on y (as a rule, it is better to do induction on the variable farthest to the right in an expression you are going to work with, because of the forms of axioms 6-9).

The basis step will be $(\forall x.x + 0 = 0 + x)$

The induction hypothesis will be $(\forall x.x + k = k + x)$ (k being an arbitrary number we introduce).

The induction goal will be $(\forall x.x + S(k) = S(k) + x)$

This gives us the following proof outline:

Goal: $(\forall xy.x + y = y + x)$ We prove this by induction on y.

Basis Goal 1: $(\forall x.x + 0 = 0 + x)$

Let k be chosen arbitrarily.

Ind Hyp 1: $(\forall x.x + k = k + x)$

Induction Goal: $(\forall x.x + S(k) = S(k) + x)$

We now proceed to fill in the complete proof (though not without further comments about what we are doing!!!) The reason we are numbering the basis and induction items is that there will be subproofs of this proof which are induction proofs themselves and have their own bases and induction steps.

Goal: $(\forall xy.x + y = y + x)$ We prove this by induction on y.

Basis Goal 1: $(\forall x.x + 0 = 0 + x)$ We prove this by induction on x!

Basis Goal 2: $0 + 0 = 0 + 0$

1: $0 + 0 = 0 + 0$ ref = [That was easy!]

Ind Hyp 2 (2): $k + 0 = 0 + k$

Ind Goal 2: $S(k) + 0 = 0 + S(k)$

3: $0 + S(k) = S(0 + k)$ ax 7 $x := 0; y := k$

4: $0 + S(k) = S(k + 0)$ subs using (2) [the ind hyp] into (3)

5: $k + 0 = k$ ax 6 $x := k$

6: $0 + S(k) = S(k)$ subs using line (5) into line (4)

7: $S(k) = S(k) + 0$ ax 6 $x := S(k)$

8: $0 + S(k) = S(k) + 0$ trans = 6,7

9: $S(k) + 0 = 0 + S(k)$ symm = 8, and we are done with the basis goal. I proved this differently than I did in class (I think) though the basic idea is the same.

Let k be chosen arbitrarily.

Ind Hyp 1 (2): $(\forall x.x + k = k + x)$ This is line 2 again because everything in an induction step uses local hypotheses and goes away. We could even call it line 1, since we are never going to refer to line 1 again, but the original line 1 has not vanished. It wouldn't do any harm to call this line 10, as long as you know that lines 2-9 above can't be used.

Induction Goal: $(\forall x.x + S(k) = S(k) + x)$ I'm going to start working on the left side of this because I can see what to do with it. I will get as far as I can and then I will see something else that I want to prove...by induction of course.

Let m be arbitrary (I'm not going to use l because it looks too much like a 1). Notice that I'm using the standard technique to deal with a universal quantifier instead of induction. Sometimes it works!

Goal: $m + S(k) = S(k) + m$

3: $m + S(k) = S(m + k)$ ax 7 $x := m, y := k$ working on left side as I said, because axiom 7 applies.

4: $m + S(k) = S(k + m)$ subs using (2) (the ind hyp) into (3).

Goal: I see that I need to prove $S(k) + m = S(k + m)$ to complete the proof. I prove this as a Lemma below, by induction: actually the Lemma is

$$(\forall xy.S(x) + y = S(x + y)),$$

65

a statement which looks rather like axiom 7 but isn't. The proof of the Lemma is given below; we proceed with the main proof assuming that we have it.

5: $S(k)+m = S(k+m)$ Lemma proved below, $x := k, y := m$

6: $m + S(k) = S(k + m)$ symm trans = lines 4 and 5

This completes the proof of the main theorem, once we prove the Lemma, whose free-standing proof follows. Of course we cannot use commutativity of addition in the proof of the Lemma!

Lemma: $(\forall xy.S(x) + y = S(x + y))$ Im going to prove this by induction on y; first I'm going to use the usual strategy for a universal quantifier to get rid of x.

Let a be chosen arbitrarily.

Goal: $(\forall y.S(a)+y = S(a+y))$ *This* is what we will prove by induction on y.

Basis Goal: $S(a) + 0 = S(a + 0)$

1: $S(a) + 0 = S(a)$ ax 6 $x := S(a)$

2: $a + 0 = a$ ax 6 $x := a$

3: $S(a + 0) = S(a)$ both sides line 2

4: $S(a) + 0 = S(a + 0)$ symm trans = 1,3

Let k be arbitrary

Ind Hyp (5): $S(a) + k = S(a + k)$

Ind Goal: $S(a)+S(k) = S(a+S(k))$ Notice that both sides offer opportunities to calculate using axiom 7.

6: $S(a) + S(k) = S(S(a) + k)$ ax 7, $x := S(a), y := k$ You should notice the opportunity to rewrite using the inductive hypothesis!

7: $S(a) + S(k) = S(S(a + k))$ subs into line 6 using line 5 (the ind hyp)

8: $a + S(k) = S(a + k)$ ax 7 $x := a, y := k$

9: $S(a)+S(k) = S(a+S(k))$ subs using (8) into (7) [an equation can be used to substitute in either order].

That completes the proof of the lemma and the theorem.

5.7 Right distributivity of multiplication over addition

Theorem: $(\forall xyz.x(y+z) = xy + xz)$. We strip off the first two quantifiers the easy way then prove the theorem by induction on z.

Let a, b be arbitrary.

Goal: $(\forall z.a(b+z) = ab + az)$. We prove this by induction on z.

Basis Goal: $a(b+0) = ab + a0$

1 $b + 0 = b$ ax 6 $x := b$

2 $a(b+0) = ab$ both sides line 1

3 $ab + 0 = ab$ ax 6 $x := ab$

4 $a0 = 0$ ax 8 $x := a$

5 $ab + a0 = ab$ subs into line 3 using line 4

6 basis goal, symm trans = lines 2,5

Let k be arbitrarily chosen

Induction Hypothesis: $a(b+k) = ab + ak$

Induction Goal (line 0): $a(b + S(k)) = ab + aS(k)$

1 $b + S(k) = S(b+k)$ ax 7 $x := b; y := k$

2 $a(b + S(k)) = a(S(b+k))$ both sides line 1

3 $aS(b+k) = a(b+k) + a$ ax 9 $x := a; y := b+k$ notice I kept parentheses

4 $a(b + S(k)) = a(b+k) + a$ trans = 2,3

5 $aS(k) = ak + a$ ax 9 $x := a; y := k$

6 $ab + aS(k) = ab + (ak + a)$ both sides line 5

7 $ab + (ak + a) = (ab + ak) + a$ associative law of addition

8 $ab + aS(k) = (ab + ak) + a$ trans = 6,7

9 $ab + aS(k) = a(b+k) + a$ subs 8 using line 0, ind hyp

10 Induction goal, symm trans = lines 4,9

5.8 Homework assigned 9/25/2019

A draft of this will be due on 10/4/2019 by noon (drafts accepted later but opportunities for feedback will become more limited). You should retain a photocopy of your paper when you hand it in on this date; I will give instructions on rewriting it, very possibly by email as I mark it, and a final

copy will be due on a later date. You are welcome to hand in your draft earlier than the 10/4 date (and welcome to do so electronically): I'll get started on the process earlier for those who turn in their drafts early. There will be a chance to work with others on these proof in class on 10/2.

1. Prove the following statements in formal arithmetic. These proofs should be relatively easy.2

   (a) $2 \cdot 2 = 4$
   (b) for every natural number x, $1x = x$ (recall that 1 is defined as $S(0)$; you are allowed to use the theorem $S(x) = x + 1$).

2. Using the axioms stated in the notes and the proof style in the notes and my board examples, prove the listed statements (this doesn't mean anything different from "prove the following statements in formal arithmetic" in the first part: it is a clarification). You may need to state and prove some lemmas. You may use theorems that you have already proved (you do not need to prove them in the exact order given). Make sure you show all parentheses in arithmetic expressions; you need to make uses of associativity explicit.

   In proving each of these theorems, you may freely use earlier theorems on this list, whether you have proved them already or not. You may not use a later one to prove an earlier one. You will get a satisfactory grade if you complete **two** of these successfully.

   I may add some hints about lemmas needed or the optimal order in which to do the problems.

   the associative law of addition no special hints for this one, it looks pretty direct.

   the left distributive property of multiplication over addition (of course if you can prove commutativity of multiplication first this will be easy [but I do not recommend this: I think the order in which I give the theorems is the correct one to prove them in]): $(x + y)z = xz + yz$.
   This is in no way an easy consequence of the right distributive law already proved, without knowing that multiplication is commutative!

looks clean and routine. Don't forget that you have to explicitly show uses of commutativity and associativity of addition.

the associative property of multiplication I think this comes out very neatly if you use previously proved results. I think the order in which I give the theorems is the correct one.

the commutative property of multiplication This is the only one where I really feel you need hints.

You need to prove $0 \cdot x = 0$ (by induction of course) and that will sew up the basis step.

Given the things already proved, the neatest lemma to prove for the induction step might be $1 \cdot x = x$ (which I assign above!). There is a lemma you can prove which is more like the one in the addiition proof, but since you have left distributivity, $1 \cdot x = x$ should do it. The lemma you would otherwise need to prove is $S(x) \cdot y = x \cdot y + y$ (and no, this does not follow from left distributivity unless you prove $1 \cdot x = x$).

5.9   Cancellation law of addition; definition of subtraction

In these last few sections we make use of axioms 3 and 4, which we have not used much so far, and use them to introduce subtraction and order.

Theorem: $(\forall xyz.x + z = y + z \rightarrow x = y)$. We get rid of the first two quantifiers the easy way, then prove the theorem by induction on z.

> **Let** a, b be chosen arbitrarily.
>
> **Goal:** $(\forall z.a + z = b + z \rightarrow a = b)$. We prove this by induction.
>
> > **Basis Goal:** $a + 0 = b + 0 \rightarrow a = b$
> >
> > > **Assume (1):** $a + 0 = b + 0$
> > >
> > > **Goal:** $a = b$
> > >
> > > **2** $a + 0 = a$ ax 6 $x := a$
> > >
> > > **3** $b + 0 = b$ ax 6 $x := b$
> > >
> > > **4** $a = b$ subs into 1 (ind hyp) using lines 2,3
> >
> > **Let** k be chosen arbitrarily.

Induction Hypothesis (1): $a + k = b + k \rightarrow a = b$

Induction Goal: $a + S(k) = b + S(k) \rightarrow a = b$

 Assume(2): $a + S(k) = b + S(k)$

 Goal: $a = b$

 3 $a + S(k) = S(a + k)$ ax 7 $x := a; b := k$

 4 $b + S(k) = S(b + k)$ ax 7 $x := b; y := k$

 5 $S(a + k) = S(b + k)$ subs into 2 using lines 3,4

 6 $S(a+k) = S(b+k) \rightarrow a+k = b+k$ ax 4 $x := a+k; y := b+k$

 7 $a + k = b + k$ mp 5,6

 8 $a = b$ mp 1 (ind hyp),7

The cancellation property of addition allows us to define subtraction.

Definition: If a, b are natural numbers, we define $a - b$ as the unique natural number k such that $k + b = a$, if there is one (of course, since we know that addition is commutative, it is equivalent to say that it is the unique k such that $a + k = b$).

$a - b$ is not always defined (we are in the realm of elementary school arithmetic where we cannot subtract 3 from 1). We define a statement like "$a - b$ does not exist" or "$a - b$ is not defined" (which is not really about a number $a - b$) as meaning "$\neg(\exists k. k + b = a)$".

Observation: The fact that there is just one such k follows from the cancellation property of addition: suppose that $k + b = a$ and $k' + b = a$ (so both k and k' are candidates to be $a - b$). Then $k + b = k' + b$, so $k = k'$ by the cancellation property of addition.

Observation: In the arithmetic of real numbers, we define $\frac{a}{b}$ as the unique real number x such that $bx = a$ and we say it does not exist not only when there is no such x (when $b = 0$ and $a \neq 0$) but also when there are many such x's (when $a = b = 0$).

5.10 Definition of order relations; transitivity of less-than

We define the relation $x < y$ and prove its major properties (or at least suggest that they can be proved). This might be surprising, as it is not obvi-

ous that we have the tools to discuss the order relations without additional primitive notions. But we do!

Definition: We define $a < b$ as meaning $(\exists x. a + S(x) = b)$.

Equivalent forms: $(\exists x. a + x = b \wedge x \neq 0)$ is equivalent; remember we have proved that each natural number is either 0 or a successor, and axiom 3 tells us that no number is both. Another equivalent is "$b - a$ exists and $b - a \neq 0$".

Additional Definitions: $a > b$ is defined as $b < a$. $a \leq b$ is defined as $a < b \vee a = b$. $a \geq b$ is defined as $b \leq a$.

Theorem: $(\forall xyz. x < y \wedge y < z \rightarrow x < z)$

   Let a, b, c be arbitrarily chosen.

   Goal: $a < b \wedge b < c \rightarrow a < c$

      Assume (1): $a < b \wedge b < c$

      Goal: $a < c$

      Rewritten Goal: $(\exists x. a + S(x) = c)$

      Find: k such that $a + S(k) = c$ (this is an imperative form of our goal statement)

      2 $a < b$ from (1): rewrite as $(\exists x. a + S(x) = b)$ using def ¡

      3 $b < c$ from (1): rewrite as $(\exists x. a + S(x) = b)$ using def ¡

      Let m be a witness to 2

      4 $a + S(m) = b$ witness to line 2

      Let n be a witness to line 3

      5 $b + S(n) = c$

      6 $(a + S(m)) + S(n) = c$ subs line 4 using line 5

      7 $(a + S(m)) + S(n) = a + (S(m) + S(n))$ assoc +

      8 $S(m) + S(n) = S(m + S(n))$ ax 7 $x := S(m); y := n$

      9 $a + (S(m) + S(n)) = c$ 7,6 trans $=$

      10 $a + S(m + S(n)) = c$ subs 9 using 8

      Let (9) $k = m + S(n)$

      10 $a + k = c$ subs 10 using 9; this is what we want.

5.11 Trichotomy

An important property of order is the trichotomy principle.

Trichotomy principle: For any natural numbers a, b exactly one of the following statements is true:

    1. $a < b$

    2. $a = b$

    3. $b < a$

Trichotomy principle in our logical notation:

$$(\forall xy.(x < y \lor x = y \lor y < x)$$

$$\land \neg(x < y \land x = y) \land \neg(y < x \land x = y) \land \neg(x < y \land y < x))$$

that is, one of the three things is true and no two of them are true at the same time.

We will prove this in parts.

Theorem: $(\forall xy.x < y \lor (x = y \lor y < x))$

> **Let** a be chosen arbitrarily.
> **Goal:** $(\forall y.a < y \lor a = y \lor y < a)$
>> **Basis Goal:** $a < 0 \lor (a = 0 \lor 0 < a)$
>> **More specific goal:** $a = 0 \lor 0 < a$
>> **1** $a = 0 \lor (\exists x.S(x) = a)$ theorem proved earlier
>>> **Case 1 (from 1); Assume(1):** $a = 0$
>>> **Goal:** $a = 0 \lor 0 < a$
>>> **2** $a = 0 \lor 0 < a$ weakening line 1
>>> **Case 2 (from 1); Assume(1):** $(\exists x.S(x) = a)$
>>> **Goal:** $a = 0 \lor 0 < a$
>>> **Specified Goal:** $0 < a$, rewrite as $(\exists x.0 + S(x) = a)$
>>> **Find** n such that $0 + S(n) = a$
>>> **Let** m witness line 1

2 $S(m) = a$ witness to line 1

3 $0 + S(m) = S(m)$ ax 8 and comm $+$

4 $0 + S(m) = a$ 3,2 trans $=$

Let (5) $n = m$

6 $0 + S(n) = a$ subs into 4 using 5; which is what we want.

Let k be chosen arbitrarily

Induction Hypothesis (1): $a < k \vee (a = k \vee k < a)$ we will make three cases from this triple disjunction rather than two cases with one having two subcases

Induction Goal: $a < S(k) \vee (a = S(k) \vee S(k) < a)$

Case 1 from (1): Assume(2): $a < k$ rewrite as

$$(\exists x. a + S(x) = k)$$

Specified Goal: $a < S(k)$ rewrite as $(\exists x. a + S(x) = S(k))$

Find q such that $a + S(q) = S(k)$

Let p witness line 2

3 $a + S(p) = k$ witness line 2

4 $S(a + S(p)) = S(k)$ both sides line 3

5 $a + S(S(p)) = S(k)$ ax 7 $x := a; y := S(p)$

Let (6) $q = S(p)$

7 $a + S(q) = S(p)$ subs into 5 using 6; which is what we want.

Case 2 from (1): Assume(2): $a = k$

Specified Goal: $a < S(k)$ rewrite as $(\exists x. a + S(x) = S(k))$

Find q such that $a + S(q) = S(k)$

3 $a + 0 = a$ ax 6 $x := a$

4 $a + S(0) = S(a + 0)$ ax 7 $x := a; y := 0$

5 $a + S(0) = S(a)$ subs using 3 into 4

6 $a + S(0) = S(k)$ subs using 2 into 5

Let(7) $q = 0$

8 $a + S(q) = k$ subs using 7 into 6; this is what we want.

Case 3 from (1): Assume(2): $k < a$ rewrite as

$$(\exists x. k + S(x) = a)$$

Specified Goal: $a = S(k) \lor S(k) < a$

Let r be a witness to 2

3 $k + S(r) = a$ witness to 2

4 $S(k + S(r)) = S(a)$ both sides 3

5 $S(k+S(r)) = S(k)+S(r)$ by lemma from the comm+ proof

6 $S(k) + S(r) = S(a)$ subs using 5 into 4

7 $S(k) + S(r) = S(S(k) + r)$ ax 7 $x := S(k); y := r$

8 $S(S(k) + r) = S(a)$ subs using 7 into 6

9 $S(S(k)+r) = S(a) \to S(k)+r = a$ ax 4 $x := S(k)+r; y := a$

10 $S(k) + r = a$ mp 8,9

11 $r = 0 \lor (\exists x.S(x) = r)$

> **Case 1 from line 11 (12):** $r = 0$
>
> **Specified Goal:** $a = S(k)$
>
> **13** $S(k) + 0 = a$ subs into 10 using 12
>
> **14** $S(k) + 0 = S(k)$ ax 6 $x := S(k)$
>
> **15** $a = S(k)$ symm trans $= 13,14$; goal of case 1
>
> **Case 2 from line 11 (12):** $(\exists x.S(x) = r)$
>
> **Specified Goal:** $S(k) < a$ rewrite as
>
> $$(\exists x.S(k) + S(x) = a$$
>
> **Find** t such that $S(k) + S(t) = a$
>
> **Let** s witness 12
>
> **13** $S(s) = r$ witness to 12
>
> **14** $S(k) + S(s) = a$ subs into 10 using 14
>
> **Let (15)** $t = s$
>
> **16** $S(k) + S(t) = a$ subs into 14 using 15; goal of case 2

This completes the proof that one of the three alternatives must be true.

Lemma: $(\forall x.\neg x < x)$

> **Let** a be arbitrary
>
> **Goal:** $\neg a < a$

Assume(1): $a < a$ rewrite as $(\exists x. a + S(x) = a)$

Goal: $\bot$

Let b witness line 1

2 $a + S(b) = a$ witness to line 1

3 $a + 0 = a$ ax 6 $x := a$

4 $a + S(b) = a + 0$ subs into line 2 using line 3

5 $a + S(b) = a + 0 \rightarrow S(b) = 0$ cancellation property of $+$

6 $S(b) = 0$ mp 4,5

7 $\neg S(b) = 0$ axiom 3 x:=b

8 $\bot$ contradiction 6,7

Theorem: $(\forall x. \neg(x < y \wedge x = y))$ Notice that this actually rules out two of the cases, as clearly it also implies $(\forall x. \neg(y < x \wedge x = y))$.

Let a, b be arbitrarily chosen.

Goal: $\neg(a < b \wedge a = b)$

    Assume(1): $a < b \wedge a = b$

    Goal: $\bot$

    2 $a < b$ from 1

    3 $a = b$ from 1

    4 $a < a$ subs into 2 using 3

    5 $\neg a < a$ lemma $x := a$

    6 $\bot$ contradiction 4,5

Theorem: $(\forall xy. \neg(x < y \wedge y < x))$ This will complete the proof of trichotomy.

Let a, b be arbitrarily chosen.

    Goal: $\neg(a < b \wedge b < a)$

    Assume(1): $a < b \wedge b < a$

    Goal: $\bot$

    2 $a < b \wedge b < a \rightarrow a < a$ transitivity theorem

    3 $a < a$ mp 1,2

    4 $\neg a < a$ lemma

    5 $\bot$ contradiction 3,4

5.12 The well-ordering principle or Least Number Principle

[NOTE: I'll put notes here]

6 Real Numbers

Here I am going to write notes on what I say from Spivak, where I think it is a good idea.

Our level of rigor is different, and we for the most part *cannot* use the proof techniques of the previous section. We are working in a different domain with different rules. A statement like (P4): $(\forall xy.x + y = y + x)$ in Spivak does not mean the same thing as our theorem $(\forall xy.x + y = y + x)$ in the previous section! The theorem here means $(\forall xy \in \mathbb{R}.x + y = y + x)$, while the theorem in the previous section means $(\forall xy \in \mathbb{N}.x + y = y + x)$. Now of course the set of natural numbers $\mathbb{N}$ is a subset of the set of real numbers $\mathbb{R}$, so if we are proving something about the natural numbers in $\mathbb{R}$, we will be able to use mathematical induction.

We will be less formal in our way or writing proofs: you can look at the chapter for samples of reasoning, and I am also going to include proofs in a style I like in my notes on chapter 1.

6.1 Spivak chapter 1 notes

We are in a new theory. Our objects are called *numbers* (by which we intend the real numbers). We have as undefined notions specific real numbers 0 and 1, for any real number a a real number $-a$, for any real number a other than 0 a real number a^{-1}, and for any real numbers a and b the real numbers $a + b$ and ab.

We begin with the following set of algebraic axioms (and we do *not* assume anything from the previous section):

(P1): For any numbers x, y, z, $x + (y + z) = (x + y) + z$. [associative property of addition]

(P2): For any number x, $x + 0 = 0 + x = x$ [identity property of addition]

(P3): For any number x, $x + (-x) = (-x) + x = 0$ [inverse property of addition]

(P4): For any numbers x, y, $x+y = y+x$ [commutative property of addition]

(P5): For any numbers x, y, z, $x(yz) = (xy)z$. [associative property of multiplication]

(P6): For any number x, $x1 = 1x = x$. $1 \neq 0$ [identity property of multiplication]

(P7): For any number $x \neq 0$, $xx^{-1} = x^{-1}x = 1$ [inverse property of multiplication]

(P8): For any numbers x, y, $xy = yx$ [commutative property of multiplication]

(P9): For any numbers x, y, z, $x(y + z) = xy + xz$ [distributive property of multiplication over addition]

You have seen this litany before, and these properties have familiar names which you are free to use, which are supplied in brackets.

The order of presentation is interesting. There is a reason why P2 and P3 are presented with what look like special cases of P4 included (and similarly for P6-8). The reason is that certain basic properties of identity and inverse can be proved without P4 (or P8 in the case of multiplication) and that we might like to know this because we sometimes work in other systems where commutativity does not hold in general but analogues of these results do hold. For example in the theory of matrices we have $IA = AI = A$ and $AA^{-1} = A^{-1}A = I$, but not in general $AB = BA$.

Here are some theorems which follow from $P1 - 3$ without use of P4.

Theorem (T1): For any a, x, if $a + x = a$ then $x = 0$.

Comment: This implies that there is only one identity element 0, but it is an even stronger statement; it says that no x can act like the identity for even a single a without being the identity.

Proof: Assume $a + x = a$. Then $(-a) + (a + x) = (-a) + a$ (doing the same thing to two equal expressions). By P1, $(-a + a) + x = (-a) + (a + x)$ so by transitivity $(-a + a) + x = (-a) + a$. Applying P3 twice, we get $0 + x = 0$. $x = 0 + x$ by P2, so by transitivity of equality we get $x = 0$.

That is a proof. Here is an acceptable proof in a line format with justifications.

77

Goal: $a + x = a \rightarrow x = 0$

Assume (1) $a + x = a$

Goal $x = 0$

2 $(-a) + (a + x) = (-a) + a$ did same thing to both sides of line 1

3 $((-a) + a) + x = (-a) + a$ P1 (notice I simply allowed a substitution justified by P1)

4 $0 + x = 0$ P3 twice (once again, I made subs justified by P3 directly)

5 $x = 0$ by P2 which is our goal.

or

Assume $a + x = a$. Then $0 = [P3](-a) + a = [\text{assumption}]$ $(-a) + (a + x) = [P1]((-a) + a) + x = [P3]0 + x = [P2]x$ establishing that $0 = x$, or equivalently $x = 0$.

I am exhibiting a variety of styles here deliberately. We are more relaxed that in the last section.

Theorem (T1b): If $x + a = a$ then $x = 0$.

Proof: Exactly as above, but add $-a$ to both sides on the right.

Theorem (T2): If $a + x = 0$, then $x = -a$.

Proof: Assume $a + x = 0$. Then $x = [P2]0 + x = [P3]((-a) + a) + x = [P1](-a) + (a + x) = [\text{assumption}](-a) + 0 = [P2](-a)$. So $x = -a$.

For a perhaps more familiar looking proof, try adding $-a$ on the left to both sides and calculating using the axioms. But you may also notice that the idea of my chain of equations is basically the same.

Theorem (T2b): If $x + a = 0$, then $x = -a$

Proof: You should be able to do it.

Theorem (T3): If $a + x = a + y$ then $x = y$. Also, if $x + a = y + a$ then $x = y$ (cancellation property of addition)

Proof: Add $-a$ on the left (or right) to both sides. Try writing it out.

Very similar theorems can be proved using P6-8 about multiplication, 1, and multiplicative inverse (with some care about 0: the cancellation property for multiplication says that if $ax = ay$ and $a \neq 0$, then $x = y$ (this is in your homework).

Now we will prove some very familiar theorems relating the additive inverse to multiplication.

Theorem (T4): $-(-a) = a$

Proof: $-a + a = 0$ by P3. But we showed above that for any x, if $-a + x = 0$ then x has to be $-(-a)$. So $a = -(-a)$. [This requires to be read carefully: understanding some very simple proofs is like fighting your way out of a wet paper bag...]

Try writing a chain of equations, each step justified by a basic property, which starts with $-(-a)$ and ends with a.

Theorem (T5): $a0 = 0a = 0$

Proof: $ab = [P2]a(b + 0) = [P9]ab + a0$. Since $ab + a0 = ab$, it follows by the first theorem we proved that $a0 = 0$. P8 lets us deduce $0a = 0$ too.

Comment: Proving simple results of this kind should give us confidence that we have enough axioms to do everything we usually do in algebra.

Theorem (T6): $a(-b) = -(ab)$ (T6a); $(-a)b = -(ab)$ (T6b); $(-a)(-b) = ab$ (T6c)

Proof: $ab + a(-b) = [P9]a(b + (-b)) = [P3]a0 = $ [theorem just proved]0; since $ab + a(-b) = 0$, and we have shown that $ab + x = 0$ implies $x = -(ab)$, we have $a(-b) = -(ab)$.

We also have $(-a)b = [P8]b(-a) = $[result just proved]$-(ba) = [P8] -(ab)$, so $(-a)b = -(ab)$

and $(-a)(-b) = $ [first part of this theorem] $- ((-a)b) = $ [second part of this theorem] $- (-(ab)) = $ [earlier theorem]ab, so $(-a)(-b) = ab$.

You could describe the last result as laws of signs, but note that the last part for example does *not* say "the product of two negative numbers is a positive number". There is no reason to believe that a number $-x$ is negative (*every* number can be written in the form $-x$!). We haven't even said what

a negative number is, and worse yet there are systems which satisfy all the axioms given so far in which we cannot even say what a positive or negative number is.

I think that the same philosophy which motivated putting instances of commutativity in P2, P3, P6, P7 could have motivated giving $(x + y)z = xz + yz$ as another statement in P9, but Spivak proves this using P8 and P9:

Theorem (P9'): $(x + y)z = xz + yz$

Proof: $(x + y)z = [P8]z(x + y) = [P9]zx + zy = [P8 \text{ twice}]xz + yz.$

Notice that in the theory of matrices where we do not have $AB = BA$ we do have both left and right distributivity of matrix multiplication over matrix addition.

Now we can for example verify the FOIL method (You may not use this as a line justification in homework 3):

Theorem (FOIL): $(x + y)(z + w) = xz + xw + yz + yw$

Proof: $(x+y)(z+w) = [P9']x(z+w)+y(z+w) = [P9 \text{ twice}]xz+xw+yz+yw$

Omitting parentheses is of course justified by P1.

Here is a very important theorem.

Zero Factor Theorem: If $ab = 0$ then either $a = 0$ or $b = 0$.

Proof: Goal: $(\forall xy.xy = 0 \rightarrow x = 0 \vee y = 0)$

$\qquad$ **Let** a, b be arbitrarily chosen numbers.
$\qquad$ **Goal:** $ab = 0 \rightarrow (a = 0 \vee b = 0)$
$\qquad\qquad$ **Assume(1)** $ab = 0$
$\qquad\qquad$ **Goal:** $a = 0 \vee b = 0$
$\qquad\qquad\qquad$ **Assume(2):** $a \neq 0$
$\qquad\qquad\qquad$ **Goal:** $b = 0$
$\qquad\qquad\qquad$ **3** $a^{-1}(ab) = a^{-1}0$ both sides 1
$\qquad\qquad\qquad$ **4** $(a^{-1}a)b = 0$ P5 on left, multiplication by 0 theorem on right
$\qquad\qquad\qquad$ **5** $1b = 0$ P7 and line 2 (you need $a \neq 0$!)
$\qquad\qquad\qquad$ **6** $b = 0$ P6

Notice the effectiveness of our formal logical strategies from previous sections! Do note that this is actually a biconditional: by T5 above, if $a = 0$ or $b = 0$ then $ab = 0$.

For future reference, we give some definitions.

Definition (def sub): $x - y$ is defined as $x + (-y)$.

Definition (def div): $\frac{x}{y}$ is defined as xy^{-1}, when $y \neq 0$.

The properties P1 through P9 are not enough to specify the real number system. They are not enough to prove $1 + 1 \neq 0$!

We all know the facts summarized in the following tables.

+	even	odd
even	even	odd
odd	odd	even

*	even	odd
even	even	even
odd	even	odd

For example, the sum of an odd integer and an odd integer is an even integer. The product of an even integer and an odd integer is an even integer.

The properties P1,P4,P5,P8,P9 obviously hold because they hold for integers. `even` is the additive identity 0 here (because 0 is even); `odd` is the multiplicative identity 1. Notice that each of the two "numbers" is its own additive inverse, so P3 holds; this is not surprising because P3 holds in the integers. P7 does not hold in the integers, but it does hold here: $1 = $ `odd` is its own multiplicative inverse.

So if we replace `even` with 0, `odd` with 1 in the tables above, we get a system which satisfies all the axioms P1-P9, and also satisfies the rather surprising $1 + 1 = 0$. What does this mean? It means that our axioms cannot disprove $1 + 1 = 0$! We need more axioms for a full description of the real numbers.

In class I gave the more complicated example of mod 5 arithmetic; in general, mod p arithmetic for any prime p (the arithmetic on remainders on division by p) will give a system which satisfies axioms P1-P9. If p is not prime, P7 will fail but the other axioms will hold. The other axioms hold fairly obviously because of the way modular arithmetic is built on top of the integers, which have all the properties except P7. The fact that P7 holds

when the modulus is prime is a surprise which requires a little number theory to demonstrate.

The additional properties we need can be described as properties of order, though they are not presented exactly that way. We introduce a new primitive notion, a set P of numbers called the "positive" numbers.

P10 For each number x, exactly one of the following is true: $x \in P$; $x = 0$; $-x \in P$.

P11: For any numbers x,y, if $x \in P$ and $y \in P$, then $x + y \in P$.

P11: For any numbers x,y, if $x \in P$ and $y \in P$, then $xy \in P$.

Now we prove that $1 + 1 \neq 0$ (if you were worried).

Theorem: $1 \in P$

Proof: By P10, exactly one of the following is true, $1 \in P$, $1 = 0$, and $-1 \in P$. $1 = 0$ is ruled out by P6. Suppose $-1 \in P$; it follows by P10 that $-(-1) = 1 \notin P$ (because only one of $-1 \in P$, $-1 = 0$, $-(-1) \in P$ can be true). But it also follows by P12 that $(-1)(-1) = 1 \in P$, which is a contradiction. Since both other alternatives are ruled out, $1 \in P$.

Theorem: $1 + 1 \neq 0$

Proof: $1 \in P$, so $1 + 1 \in P$ by P11, so $1 + 1 \neq 0$ by P10 applied to $1 + 1$.

We were told these were properties of order, but no notion of order has yet been introduced.

Definition $x < y$ is defined as meaning $y - x \in P$. $y > x$ is defined as meaning $x < y$. $x \leq y$ is defined as meaning $x < y \vee x = y$. $y \geq x$ is defined as meaning $x \leq y$.

In problem 8, a different approach is indicated. Suppose instead that $x < y$ is our basic notation and we define $x \in P$ as meaning $x > 0$ and adopt the following axioms:

P10' For any numbers x, y, exactly one of the following is true: $x < y; x = y; x > y$

82

P11' For any numbers x, y, z, if $x < y$ and $y < z$, then $x < z$.

P12' For any numbers x, y, z, if $x < y$ then $x + z < y + z$.

P13' For any numbers x, y, z, if $x < y$ and $z > 0$, then $xz < yz$.

We will show here that we can prove all of these statements starting with P10-12. In problem 8 you are asked to prove P10-12 starting with P10'-P13'.

Lemma (T2c): $b - a = 0$ iff $b = a$. Assume $b = a$; then $b - a = a - a = a + (-a) = 0$. Assume $b - a = 0$: then $b = b + (-a) + a = 0 + a = a$.

Lemma (add inv sub): $a - b = -(b - a)$. $(a - b) + (b - a) = a + -b + b + -a = a + 0 + -a = a + -a = 0$; apply T2.

Proof of P10': Let a, b be arbitarily chosen numbers. $b - a$ is a number: by P10 exactly one of the following is true:

$b - a \in P$, which is equivalent to $a < b$;

$b - a = 0$, which implies $b = a$ by T2c;

$-(b - a) \in P$, which is equivalent to $a - b \in P$, which is equivalent to $b < a$, which is equivalent to $a > b$.

So exactly one of $a < b$, $a = b$, $a > b$ is true.

Proof of P11': Assume $x < y$ and $y < z$. Thus $y - x \in P$ and $z - y \in P$. Thus by P11 $(y - x) + (z - y) \in P$. By algebra, $z - x \in P$, that is $x < z$.

Proof of P12': Assume $x < y$. This means $y - x \in P$. $y - x = (y + z) - (x + z)$ by algebra; thus $x + z < y + z$ follows from $x < y$ (in fact, it is exactly equivalent; we could state a biconditional here as easily as an implication).

Proof of P13': Assume $x < y$ and $z > 0$. This is equivalent to $y - x \in P$ and $z \in P$. By P12 we get $(y - x)z \in P$. By algebra, we get $yz - xz \in P$, that is $xz < yz$.

These proofs are somewhat briefer than what I expect from you in problem 8. Please spell out the algebra using properties. Remember that when working in the other direction, your basic undefined notion is $<$, and $x \in P$ is defined as $x > 0$, not the other way around. In problem 8 you will be using P10'-P13' to prove each of P10-P12 (with the help of P1-9).

6.2 Homework from chapter 1 in Spivak, first part assigned 10/10/2019

This is due on Monday after Test I (which is Oct 16 2019) (just three problems). These will be useful for test preparation.

Chapter 1 problems 1, 2, 3 all parts.

I will do examples of the kind of work expected on Monday in class, and I will give you a very precise idea of which proofs to be prepared to write for the exam Wednesday. In general, there will be a thorough discussion of exam readiness on Monday.

Comments on Chapter 1 and 2 Homework I know, I haven't assigned anything from chapter 2 yet. But this assignment may not be complete and will shortly be followed by another one.

In chapter 1, you are expected to use the style I use here in the notes (or one of the styles). You need to state when you are using the various properties P1-12, and show steps (with the one exception that you may regroup and reorder long sums and products and say "regrouping addition" or "many applications of P1 and P4" or "regrouping multiplication" or "many applications of P5 and P8", as appropriate.) In chapter 2 you may assume basic knowledge from algebra, but use common sense based on what is being proved; the statement being proved may indicate that certain things which in other contexts you could assume should be spelled out.

I have given names to various theorems in this draft; currently I haven't used these names in line references consistently, but I'm planning to go back and fix that. You may use these names in your homework.

6.3 Epsilon-Delta Proofs

We remind ourselves of the definition, including the important "for all x" which is often left out.

$$\lim_{x \to a} f(x) = L$$

means "for each $\epsilon > 0$ there is a $\delta > 0$ such that for any x, if $0 < |x - a| < \delta$ then $|f(x) - L| < \epsilon$.

Example 1:
$$\lim_{x \to 3} x^2 = 9$$

We outline the proof:

Choose an $\epsilon_0 > 0$ arbitrarily.

Let $\delta_0 = \ldots$

Choose x arbitrarily.

Assume that $0 < |x - 3| < \delta_0$

Goal: $|x^2 - 9| < \epsilon_0$

The next phase is scratch work to figure out what δ_0 should be. r We aim to make $|x^2 - 9| < \epsilon_0$.

$|x^2 - 9| = |x + 3||x - 3| < \epsilon_0$ will be true if $|x - 3| < \frac{\epsilon_0}{|x+3|}$

We cannot set $\delta_0 = \frac{\epsilon_0}{|x+3|}$, because this expression depends on x. What we need is $|x - 3| < \frac{\epsilon_0}{???} < \frac{\epsilon_0}{|x+3|}$ where ???, whatever it is, does not depend on x. We need ??? to be *greater than* $|x + 3|$ (so that its reciprocal will be smaller). To get an upper bound on $|x + 3|$, we impose an upper bound on x: the only way we have to do this is to make stipulations about δ_0. If we impose $\delta_0 \leq 1$, then we get $|x-3| < 1$, which is equivalent to $2 < x < 4$. We then get $5 < x + 3 < 7$, and since $x + 3 > 5 > 0$ we have $|x + 3| = x + 3 < 7$. 7 is the desired upper bound. So we get $|x - 3| < \frac{\epsilon_0}{7} < \frac{\epsilon_0}{|x+3|}$ implies $|x^2 - 9| < \epsilon_0$ as long as we also stipulated $|x - 3| < 1$. So a workable value of δ_0 is $\min(1, \frac{\epsilon_0}{7})$.

Now we continue the proof, setting $\delta_0 = \min(1, \frac{\epsilon_0}{7})$.

Since $|x - 3| < \delta_0$, we also have $|x - 3| < 1$ and $|x - 3| < \frac{\epsilon_0}{7}$. From this we deduce $5 < |x + 3| = x + 3 < 7$ just as we did above in the scratch work. Now $|x^2 - 9| = |x + 3||x - 3| < 7|x - 3| < 7\frac{\epsilon_0}{7} = \epsilon_0$.

Example 2: $\lim_{x \to 0} \frac{|x|}{x}$ does not exist.

We prove this by assuming that $\lim_{x \to 0} \frac{|x|}{x} = L$ for some L and deducing a contradiction.

Since $\lim_{x \to 0} \frac{|x|}{x} = L$, there is a δ_0 such that for any x, if $0 < |x - 0| = |x| < \delta_0$ then $|\frac{|x|}{x} - L| < 1$ (here setting $\epsilon = 1$), Now set $x = \frac{\delta_0}{2}$: we get $|1 - L| < 1$, and similarly if we set $x = -\frac{\delta_0}{2}$, we get $|-1 - L| = |L + 1| < 1$

So we get $2 = |(1 - L) + (1 + L)| \le |1 - L| + |1 + L| < 1 + 1 = 2$, a contradiction.

There are definitely details of motivation for you to track down and fill in here.

We prove some actual theorems.

Theorem (limit of a constant):

$$\lim_{x \to a} c = c$$

Notice that the two c's mean different things. The first one is the constant function whose value is c, the second is the real number c.

Choose $\epsilon_0 > 0$. Let $\delta_0 = 1$. Choose x arbitrarily such that $0 < |x - a| < \delta_0 = 1$. Our goal is to show that $|c - c| < \epsilon_0$. But $|c - c| = 0 < \epsilon_0$ is definitely true! In fact it simply does't matter what δ_0 we choose.

Theorem (limit of the identity function):

$$\lim_{x \to a} x = a$$

Choose $\epsilon_0 > 0$. Let $\delta_0 = \epsilon_0$. Choose x such that $0 < |x - a| < \delta_0 = \epsilon_0$. Our goal is $|x - a| < \epsilon$, but this is included in our assumption, so we are done.

Theorem (constant multiple property of limits): If $\lim_{x \to a} f(x) = L$, then $\lim_{x \to a} cf(x) = cL$.

Outline of start of proof:

Assume $\lim_{x \to a} f(x) = L$.

Choose $\epsilon_0 > 0$ arbitrarily.

Let $\delta_0 = \ldots$

Choose x arbitrarily.

Assume $0 < |x - a| < \delta_0$.

Goal: $|cf(x) - cL| < \epsilon_0$.

Now we are at the scratch work phase.

87

$|cf(x) - cL| = |c||f(x) - L| < \epsilon_0$ will be enforced if $|f(x) - L| < \frac{\epsilon_0}{|c|}$.

Because we assumed $\lim_{x \to a} f(x) = L$, we know that there is a $\delta_0 > 0$ such that for each x with $0 < |x - a| < \delta_0$, we have $|f(x) - L| < \frac{\epsilon_0}{|c|}$.

Let this be our choice of δ_0 in the proof.

The proof resumes, choosing δ_0 so that for each x with $0 < |x - a| < \delta_0$, we have $|f(x) - L| < \frac{\epsilon_0}{|c|}$.

Since we have assumed $0 < |x - a| < \delta_0$ it follows that $|f(x) - L| < \frac{\epsilon_0}{|c|}$. Thus $|cf(x) - cL| = |c||f(x) - L| < |c|\frac{\epsilon_0}{|c|} = \epsilon_0$, completing the proof.

It is important to notice that we do not state this theorem "$\lim_{x \to a} cf(x) = c \lim_{x \to a} f(x)$". Suppose that $c = 0$ and $\lim_{x \to a} f(x)$ is undefined. In this case the left side of this equation would be 0 and the right side would be undefined.

Theorem (addition property of limits): if $\lim_{x \to a} f(x) = L$ and $\lim_{x \to a} g(x) = M$ then

$$\lim_{x \to a} f(x) + g(x) = L + M.$$

Note first of all that

$$\lim_{x \to a} f(x) + g(x) = \lim_{x \to a} f(x) + \lim_{x \to a} g(x)$$

does not say the same thing. If $f(x) = \frac{1}{x}$ and $g(x) = -\frac{1}{x}$ then the left side is 0 and the right side is undefined.

The proof starts.

Assume $\lim_{x \to a} f(x) = L$ and $\lim_{x \to a} g(x) = M$.

Choose $\epsilon_0 > 0$.

Let $\delta_0 = \ldots$

Choose x arbitrarily.

Assume $0 < |x - a| < \delta_0$.

Goal: $|(f(x) + g(x)) - (L + M)| < \epsilon_0$

We pause for scratch work.

$|(f(x) + g(x)) - (L + M)| = |(f(x) - L) + (g(x) - M)| \leq |f(x) - L| + |g(x) - M|$. This will be less than ϵ_0 if we make $|f(x) - L| < \frac{\epsilon_0}{2}$

and $|g(x) - M| < \frac{\epsilon_0}{2}$ By limit assumptions, we can choose δ_1 so that if $0 < |x - a| < \delta_1$, then $|f(x) - L| < \frac{\epsilon_0}{2}$ and choose δ_2 so that if $0 < |x - a| < \delta_2$, then $|g(x) - M| < \frac{\epsilon_0}{2}$. The point is that the limit statements about f and g allow us to find δ's corresponding to any value of ϵ, in this case a value half as large as the value being considered for the sum function. Let $\delta_0 = \min(\delta_1, \delta_2)$. Continue the proof.

Since we have assumed $0 < |x - a| < \delta_0$, we also have $0 < |x - a| < \delta_1$ and $0 < |x - a| < \delta_2$, so we have $|f(x) - L| < \frac{\epsilon_0}{2}$ and $|g(x) - M| < \frac{\epsilon_0}{2}$. Thus $|(f(x) + g(x)) - (L + M)| = |(f(x) - L) + (g(x) - M)|$ $\leq |f(x) - L| + |g(x) - M| < \frac{\epsilon_0}{2} + \frac{\epsilon_0}{2} = \epsilon_0$.

Another example:

Example: If $a \neq 0$, $\lim_{x \to a} \frac{1}{x} = \frac{1}{a}$

Choose $\epsilon_0 > 0$ arbitrarily.

Let $\delta_0 = \ldots$

Choose x arbitrarily.

Assume $0 < |x - a| < \delta_0$.

Goal: $|\frac{1}{x} - \frac{1}{a}| < \epsilon_0$.

Now it is time for scratch work. $|\frac{1}{x} - \frac{1}{a}| = |\frac{x-a}{ax}|$. This will be forced to be less than ϵ_0 if $|x - a| < \epsilon_0|a||x|$. The last expression does not work as a delta because it contains x. What we need is $|x - a| < \epsilon_0|a||???| < \epsilon_0|a||x|$ where $???$ must be a lower bound on the value of $|x|$.

To bound the value of $|x|$ away from 0, we assume $|x - a| < \frac{|a|}{2}$. From this we conclude that $|x| = |a - (a - x)| \geq |a| - |a - x| > |a| - \frac{|a|}{2} = \frac{|a|}{2}$. Now this gives us $|x - a| < \epsilon_0|a||\frac{|a|}{2}|$ as a target. Choose $\delta_0 = \min(\frac{|a|}{2}, \epsilon_0\frac{|a|^2}{2})$.

Since we have assumed $0 < |x - a| < \delta_0$, we also have $0 < |x - a| < \delta_1$ and $0 < |x - a| < \delta_2$, so we have $|x - a| < \frac{|a|}{2}$ and $|x - a| < \epsilon_0|a||\frac{|a|}{2}|$. Because $|x - a| < \frac{|a|}{2}$, we have $|x| > \frac{|a|}{2}$. Now $|\frac{1}{x} - \frac{1}{a}| = \frac{|x-a|}{|ax|} < \frac{2|x-a|}{a^2} < \frac{2(\epsilon_0(\frac{|a|^2}{2}))}{a^2} = \epsilon_0$.

6.4 Homework posted March 6th 2016, due Thursday March 10th 2016

1. Prove that
$$\lim_{x \to 5} x^2 = 25$$
   directly from the definition without using limit theorems – this will be very similar to proofs done in class. For best learning results, write it out without looking at the other proofs – but it goes just the same way.

2. Prove that
$$\lim_{x \to 3} \frac{1}{x^2} = \frac{1}{9}$$
   , directly from the definition without using limit theorems.

3. (extra credit) Prove that $\lim_{x \to a} x^3 = a^3$ directly from the definition without using limit theorems. You might want to use the factoring fact $x^3 - a^3 = (x - a)(x^2 + ax + a^3)$. This is considerably harder.

4. Prove without using any of the theorems in the chapter (but of course in a way very similar to the proof of the first part of Theorem 2) that

$$\lim_{x \to a} f(x) - g(x) = L - M$$

   if

$$\lim_{x \to a} f(x) = L$$

   and

$$\lim_{x \to a} g(x) = M$$

   You will get the most benefit out of this if you first carefully read the proof of the first part of theorem 2, then set it aside and work through it directly just as I do in class, reasoning about what you need to make it work. The facts about absolute values that you use are similar, but the computational details will be different.

6.5   More Epsilon-Delta Proofs

Theorem (multiplication property of limits): If $\lim_{x \to a} f(x) = L$ and $\lim_{x \to a} g(x) = M$ then $\lim_{x \to a} f(x)g(x) = LM$.

Proof: Assume $\lim_{x \to a} f(x) = L$ and $\lim_{x \to a} g(x) = M$.

Choose $\epsilon_0 > 0$ arbitrarily.

Let $\delta_0 = \ldots$ [something we will figure out in a minute]

Choose x arbitrarily.

Assume $0 < |x - a| < \delta_0$.

Goal: $|f(x)g(x) - LM| < \epsilon_0$.

Intermission for Scratch Work: We need to figure out what δ_0 needs to be.

We want to make $|f(x)g(x) - LM| < \epsilon_0$ by controlling $|f(x) - L|$ and $|g(x) - M|$, which we can do using the assumptions about existence of limits. To do this we need to turn this into a form we can express in terms of $|f(x) - L|$ and $|g(x) - M|$.

$$|f(x)g(x) - LM| = |f(x)g(x) - f(x)M + f(x)M - LM| \leq$$

[now use triangle inequality]

$$|f(x)g(x) - f(x)M| + |f(x)M - LM| = |f(x)||g(x) - M| + |f(x) - L||M|.$$

Notice that we now have things in terms of $|f(x) - L|$ and $|g(x) - M|$. It might be worth checking for yourself that $Lg(x)$ works just as well.

So now we want to make $|f(x)||g(x) - M| + |f(x) - L||M| < \epsilon_0$, which we can do by making $|f(x)||g(x) - M| < \frac{\epsilon_0}{2}$ and $|f(x) - L||M| < \frac{\epsilon_0}{2}$.

The second is simpler to handle, so we will deal with it first. $|f(x) - L||M| < \frac{\epsilon_0}{2}$ looks easy to handle – divide both sides by M. But we have to take into account that M might be 0. $|f(x) - L||M| < |f(x) - L|(|M| + 1) < \frac{\epsilon_0}{2}$ can be enforced by making $|f(x) - L| < \frac{\epsilon_0}{2(|M|+1)}$. So we choose a δ_1 such that for any x, if $0 < |x - a| < \delta_1$ then $|f(x) - L| < \frac{\epsilon_0}{2(|M|+1)}$. We can do this because of the known limit of $f(x)$ at a.

Now we deal with the first. $|f(x)||g(x) - M| < \frac{\epsilon_0}{2}$ is equivalent to $|g(x) - M| < \frac{\epsilon_0}{2|f(x)|}$. This is not satisfactory because of the dependence on x. We need $|g(x) - M| < \frac{\epsilon_0}{2???} < \frac{\epsilon_0}{2|f(x)|}$, where ??? does not depend on x: what is needed is an upper bound on $|f(x)|$. If we choose a δ_2 such that for any x, if $0 < |x - a| < \delta_2$ we have $|f(x) - L| < 1$, then we further note that under these conditions we also have $|f(x)| = |(f(x) - L) + L| \le |f(x) - L| + |L| < 1 + |L|$. This gives the desired upper bound on $f(x)$, so what we want is $|g(x) - M| < \frac{\epsilon_0}{2(|L|+1)} < \frac{\epsilon_0}{2|f(x)|}$, which we can enforce by choosing a δ_3 such that for any x, if $0 < |x - a| < \delta_3$ then $|g(x) - M| < \frac{\epsilon_0}{2(|L|+1)}$.

The value for δ_0 that we want is the minimum of δ_1, δ_2 and δ_3 chosen above.

Proof resumes: Choose δ_0 as indicated in the scratch work and continue the proof.

$$|f(x)g(x) - LM| = |f(x)g(x) - f(x)M + f(x)M - LM|$$

$$\le |f(x)g(x) - f(x)M| + |f(x)M - LM| = |f(x)||g(x) - M| + |M||f(x) - L|.$$

Now we have $|f(x) - L| < \frac{\epsilon_0}{2(|M|+1)}$ because $0 < |x - a| < \delta_0 \le \delta_1$.

Thus

$$|f(x)||g(x) - M| + |M||f(x) - L| < |f(x)||g(x) - M| + |M|\frac{\epsilon_0}{2(|M| + 1)}$$

$$< |f(x)||g(x) - M| + \frac{\epsilon_0}{2}.$$

Further, we have $|f(x) - L| < 1$, because $0 < |x - a| < \delta_0 \leq \delta_2$, so $|f(x) < 1 + |L|$ (derivation above in the scratch work), and $|g(x) - M| < \frac{\epsilon_0}{2(|L|+1)}$, because $0 < |x - a| < \delta_0 \leq \delta_3$. Thus

$$|f(x)||g(x) - M| + \frac{\epsilon_0}{2} < (1 + |L|)(\frac{\epsilon_0}{2(|L| + 1)}) + \frac{\epsilon_0}{2} = \frac{\epsilon_0}{2} + \frac{\epsilon_0}{2} = \epsilon_0,$$

completing the proof.

Lemma: a variation on the triangle inequality. Notice that $|a| = |(a - b) + b| \leq |a - b| + |b|$. Subtracting $|b|$ from both sides we get $|a| - |b| \leq |a - b|$. Because $|a - b| = |b - a|$, we also get $|b| - |a| \leq |a - b|$.

Theorem (reciprocal property of limits): If $\lim_{x \to a} f(x) = L \neq 0$, then

$$\lim_{x \to a} \frac{1}{f(x)} = \frac{1}{L}.$$

Proof: Assume $\lim_{x \to a} f(x) = L \neq 0$.

Choose $\epsilon_0 > 0$ arbitrarily.

Let $\delta_0 = \ldots$ [something we will figure out in a minute]

Choose x arbitrarily.

Assume $0 < |x - a| < \delta_0$.

Goal: $|\frac{1}{f(x)} - \frac{1}{L}| < \epsilon_0$.

Intermission for scratch work: We want to make $|\frac{1}{f(x)} - \frac{1}{L}| < \epsilon_0$.

$|\frac{1}{f(x)} - \frac{1}{L}| = |\frac{L - f(x)}{Lf(x)}| = \frac{|L - f(x)|}{|L||f(x)|} = \frac{|f(x) - L|}{|L||f(x)|}$. To force $\frac{|f(x) - L|}{|L||f(x)|} < \epsilon_0$ is to force $|f(x) - L| < \epsilon_0 |L||f(x)|$, which we do by forcing $|f(x) - L| < \epsilon_0 |L|(???) < \epsilon_0 |L||f(x)|$, where ??? is a lower bound on $|f(x)|$ not depending on x. We know that $f(x)$ can be forced to be close to L. We also know that we want to keep $f(x)$ away from 0. This suggests

that we want $|f(x) - L| < \frac{|L|}{2}$, half the distance from $f(x)$ to 0. Note that if $|f(x) - L| < \frac{|L|}{2}$ then we have $|L| - |f(x)| \leq |f(x) - L| < \frac{|L|}{2}$ (by the variation on the triangle inequality), and from $|L| - |f(x)| < \frac{|L|}{2}$ we can deduce $|L| - \frac{|L|}{2} < |f(x)|$, that is $|f(x)| > \frac{|L|}{2}$. This gives us the desired lower bound on $|f(x)|$, so we want to enforce $|f(x) - L| < \epsilon_0|L|(\frac{|L|}{2}) < \epsilon_0|L||f(x)|$, so it is sufficient to choose a δ_1 such that for any x, if $0 < |x - a| < \delta_1$, we have $|f(x) - L| < \frac{|L|}{2}$ (to give the lower bound on $f(x)$) then a δ_2 such that for any x, if $0 < |x - a| < \delta_2$ we get $|f(x) - L| < \frac{\epsilon_0|L|^2}{2}$ (the bound we get when we plug in the specific lower bound on $f(x)$ for the quantity labelled ???). δ_0 can then be chosen to be the minimum of δ_1 and δ_2.

The proof resumes:
$$|\frac{1}{f(x)} - \frac{1}{L}| = \frac{|f(x) - L|}{|f(x)||L|}.$$

(algebra). Because $0 < |x - a| < \delta_0 \leq \delta_1$, we have $|f(x) - L| < \frac{|L|}{2}$, so $|f(x)| > \frac{|L|}{2}$ (shown in the scratch work), so

$$\frac{|f(x) - L|}{|f(x)||L|} < \frac{|f(x) - L|}{\frac{|L|}{2}|L|} = \frac{2|f(x) - L|}{|L|^2}.$$

Now because $0 < |x - a| < \delta_0 \leq \delta_2$, we have $|f(x) - L| < \frac{\epsilon_0|L|^2}{2}$, so

$$\frac{2|f(x) - L|}{|L|^2} < \frac{2\frac{\epsilon_0|L|^2}{2}}{|L|^2} = \epsilon_0,$$

completing the proof.

A statement and proof of the quotient property of limits is now straightforward, and does not involve any epsilons or deltas. Combine the product and reciprocal properties. Try writing out this proof!

We give some further examples. The first one should illustrate that we have general techniques for writing explicit epsilon-delta proofs for complicated functions, though once we have proved the basic theorems we really do not need to do this (it might be useful if one needed explicit error estimates for a function, though).

94

Example: Prove $\lim_{x \to 2} \frac{x}{x^2+1} = \frac{2}{5}$.

We begin with the scratch work. Choose an arbitrary ϵ_0. We want to find a δ_0 to make $\left|\frac{x}{x^2+1} - \frac{2}{5}\right| < \epsilon_0$ if we have $0 < |x - 2| < \delta_0$.

We want to bound $\left|\frac{x}{x^2+1} - \frac{2}{5}\right|$ by something we can express in terms of $|x - 2|$. We achieve this by a trick (which should look similar to what we did in proving the multiplication property of limits.

$\left|\frac{x}{x^2+1} - \frac{2}{5}\right| = \left|\frac{x}{x^2+1} - \frac{2}{x^2+1} + \frac{2}{x^2+1} - \frac{2}{5}\right| \leq \left|\frac{x}{x^2+1} - \frac{2}{x^2+1}\right| + \left|\frac{2}{x^2+1} - \frac{2}{5}\right|$. The term on the left of the sum just written clearly has $|x - 2|$ as a factor, and the term on the right does too after a little straightforward algebra.

We can make $\left|\frac{x}{x^2+1} - \frac{2}{x^2+1}\right| + \left|\frac{2}{x^2+1} - \frac{2}{5}\right| < \epsilon_0$ by making $\left|\frac{x}{x^2+1} - \frac{2}{x^2+1}\right| < \frac{\epsilon}{2}$ and $\left|\frac{2}{x^2+1} - \frac{2}{5}\right| < \frac{\epsilon}{2}$.

To make $\left|\frac{x}{x^2+1} - \frac{2}{x^2+1}\right| < \frac{\epsilon}{2}$: $\left|\frac{x}{x^2+1} - \frac{2}{x^2+1}\right| = \frac{|x-2|}{x^2+1}$. It is sufficient to make $|x - 2| < \frac{\epsilon_0(x^2+1)}{2}$. The bounding expression is not satisfactory because it depends on x: $|x - 2| < \frac{\epsilon_0(???)}{2} < \frac{\epsilon_0(x^2+1)}{2}$ is what we want, where ??? needs to be a lower bound on $x^2 + 1$. If we choose $\delta \leq 1$, then we have $1 < x < 3$, so $1^2 + 1 = 2$ is a suitable lower bound, and we get $|x - 2| < \frac{\epsilon_0(2)}{2} < \frac{\epsilon_0(x^2+1)}{2}$ as what we want, from which we see that we also want $\delta \leq \frac{\epsilon_0(2)}{2} = \epsilon_0$.

To make $\left|\frac{2}{x^2+1} - \frac{2}{5}\right| < \frac{\epsilon}{2}$: $\left|\frac{2}{x^2+1} - \frac{2}{5}\right| = \left|\frac{10}{5(x^2+1)} - \frac{2x^2+2}{5(x^2+1)}\right| = \frac{2(x^2-4|)}{5(x^2+1)} = \frac{2|x+2||x-2|}{5(x^2+1)}$. So it is sufficient to make $|x-2| < \frac{\epsilon_0}{2} \cdot \frac{5(x^2+1)}{2|x+2|}$. The bounding expression is unsatisfactory for the usual reasons. We need to replace $x^2 + 1$ with a lower bound on this expression (we already have one, namely 2); we also need to replace $|x + 2|$ with an upper bound: since we have $1 < x < 3$, we have $3 < |x + 2| = x + 2 < 5$, so 5 is a suitable upper bound, and what we want is $|x - 2| < \frac{\epsilon_0}{2} \cdot \frac{5(2)}{2(5)} < \frac{\epsilon_0}{2} \cdot \frac{5(x^2+1)}{2|x+2|}$, from which we can see that we also want $\delta < \frac{\epsilon_0}{2} \cdot \frac{5(2)}{2(5)} = \frac{\epsilon_0}{2}$.

Thus we want $\delta_0 = \min(1, \frac{\epsilon_0}{2})$.

So, let x be chosen arbitrarily. Assume that $0 < |x - 2| < \delta_0$.

Note that since $\delta_0 \leq 1$ we have $|x - 2| < 1$, that is $1 < x < 3$, from which we can conclude $x^2 + 1 > 2$ and $|x + 2| = x + 2 < 5$.

We also have $|x - 2| < \frac{\epsilon_0}{2}$.

Now we have $|\frac{x}{x^2+1} - \frac{2}{5}| = |\frac{x}{x^2+1} - \frac{2}{x^2+1} + \frac{2}{x^2+1} - \frac{2}{5}| \leq |\frac{x}{x^2+1} - \frac{2}{x^2+1}| + |\frac{2}{x^2+1} - \frac{2}{5}| = \frac{|x-2|}{x^2+1} + \frac{2|x+2||x-2|}{5(x^2+1)} < \frac{|x-2|}{2} + \frac{2(5)|x-2|}{5(2)} = \frac{3}{2}|x-2| < \frac{3}{2} \cdot \frac{\epsilon_0}{2} = \frac{3}{4}\epsilon_0 < \epsilon_0$.

Example: Prove that $\lim_{x\to 0} \frac{1}{x}$ does not exist.

Suppose it did, for the sake of a contradiction. So we would have an L such that $\lim_{x\to 0} \frac{1}{x} = L$. By this assumption (setting $\epsilon = 1$) there is a δ_0 such that if $0 < |x| < \delta_0$, we have $|\frac{1}{x} - L| < 1$, so we have $|\frac{1}{x}| = |(\frac{1}{x} - L) + L| \leq |\frac{1}{x} - L| + |L| < 1 + |L|$. This further implies that $|x| > \frac{1}{1+|L|}$. But then notice that if $0 < |x| < \min(\delta_0, \frac{1}{1+|L|})$ we have obviously that $|x| < \frac{1}{1+|L|}$, but we also have $|x| > \frac{1}{1+|L|}$ because that we have shown to follow from $0 < |x| < \delta_0$. And this is absurd.

Theorem: If $f(x) \leq g(x)$ for all x, $\lim_{x\to a} f(x) = L$ and $\lim_{x\to a} g(x) = M$, then $L \leq M$.

Suppose that $\lim_{x\to a} f(x) = L$ and $\lim_{x\to a} g(x) = M$. Suppose that $L > M$ for the sake of a contradiction.

By assumption there is δ_1 such that for any x if $0 < |x - a| < \delta_1$, then $|f(x) - L| < \frac{L-M}{2}$, from which it follows that $f(x) > L - \frac{L-M}{2} = \frac{L+M}{2}$.

By assumption there is δ_2 such that for any x if $0 < |x - a| < \delta_2$, then $|g(x) - M| < \frac{L-M}{2}$, from which it follows that $g(x) < M + \frac{L-M}{2} = \frac{L+M}{2}$.

But then it follows that for any x, if $0 < |x - a| < \min(\delta_1, \delta_2)$, it follows that $f(x) > \frac{L+M}{2} > g(x)$, so $f(x) > g(x)$ for any such x, which is a contradiction.

There is another way to prove this using the subtraction property of limits, which I will eventually add here.

Discussion of problem 8 chapter 5: Problem 8 chapter 5 raises issues which are really Math 170 issues but which are important to get right.

It is not correct to say (for example)

$$\lim_{x\to a} f(x) + g(x) = \lim_{x\to a} f(x) + \lim_{x\to a} g(x)$$

This needs an additional condition: it is only true if we know that the two limits on the right exist.

For example

$$\lim_{x \to 0} \frac{1}{x} + (-\frac{1}{x}) = \lim_{x \to 0} \frac{1}{x} + \lim_{x \to 0} -\frac{1}{x}$$

is INCORRECT. The expression on the left is equal to 0. The two limits on the right are undefined.

It is harder to get a counterexample for multiplication. The simplest one I know of is this:

$$\lim_{x \to 0} (\frac{|x|}{x})(\frac{|x|}{x}) = \lim_{x \to 0} \frac{|x|}{x} \cdot \lim_{x \to 0} \frac{|x|}{x}$$

is INCORRECT, because the limit expression on the left is equal to 1, but the two (identical) limits which appear as factors of the expression on the right are undefined.

If the functions are taken to be $\frac{|x|}{x}$ and $-\frac{|x|}{x}$, you get counterexamples for both addition and multiplication.

I have to admit being fooled by some of you:

the functions $f(x) = \ln(|x|)$ and $g(x) = \frac{1}{\ln(|x|)}$ do NOT give a multiplication counterexample. The problem is that $\lim_{x \to 0} \frac{1}{\ln(|x|)}$ exists: it is 0. So the limit of $f(x)g(x)$ at 0 exists (it is 1), the limit of $f(x)$ at 0 does not exist, but the limit of $g(x)$ at 0 does exist.

If $\lim_{x \to a} f(x)$ exists and $\lim_{x \to a} g(x)$ does not exist, then $\lim_{x \to a} f(x) + g(x)$ does not exist: assume that $\lim x \to af(x)$ exists and $\lim_{x \to a} g(x)$ does not exist and suppose for the sake of a contradiction that $\lim_{x \to a} f(x) + g(x)$ does exist: then $\lim_{x \to a} f(x) + g(x) = M$ for some M, $\lim_{x \to a} f(x) = L$ for some L by assumption, and $\lim_{x \to a} g(x) = M - L$ by the subtraction property of limits, contradicting the assumption that this limit does not exist.

For multiplication the situation is different: it is possible for $\lim_{x \to a} f(x)$ to exist, $\lim_{x \to a} g(x)$ not to exist, and $\lim_{x \to a} f(x)g(x)$ to exist. Suppose that $f(x)$ is the constant zero function (it is NOT enough to just suppose that $\lim_{x \to a} f(x) = 0$; suppose $g(x)$ is any function defined for all reals at all (such as the function equal to 1 at each rational and 0 at each irrational, which has no limit anywhere); then $\lim_{x \to a} f(x)g(x) = 0$ without any control over limits of $g(x)$ at all.

6.6 More Epsilon-Delta Proofs from Fall 2014

Things that are still missing from here will be added this weekend.

6.6.1 Quotient Rule

I will insert the full proof of the Quotient Rule that I did in class here, as an example. Be aware that I regard this as a bit too big to be a test question. The reciprocal rule, which appears above in the notes and in the book, is a better model for a test question.

Theorem: If $\lim_{x \to a} f(x) = L$ and $\lim_{x \to a} g(x) = M \neq 0$, then

$$\lim_{x \to a} \frac{f(x)}{g(x)} = \frac{L}{M}$$

Proof: This means, for each $\epsilon > 0$ there is $\delta > 0$ such that for any x, if $0 < |x - a| < \delta$ then $|f(x) - L| < \epsilon$.

Choose $\epsilon > 0$ arbitrarily.

We need to figure out what δ has to be.

Scratch work: We want $\left| \frac{f(x)}{g(x)} - \frac{L}{M} \right| < \epsilon$.

$$\left| \frac{f(x)}{g(x)} - \frac{L}{M} \right| = \left| \frac{f(x)}{g(x)} - \frac{L}{g(x)} + \frac{L}{g(x)} - \frac{L}{M} \right| \leq$$

$$\left| \frac{f(x)}{g(x)} - \frac{L}{g(x)} \right| + \left| \frac{L}{g(x)} - \frac{L}{M} \right| =$$

$$\frac{|f(x) - L|}{g(x)} + \frac{|L||g(x) - M|}{|g(x)||M|}$$

We want $\frac{|f(x) - L|}{g(x)} < \frac{\epsilon}{2}$ and $\frac{|L||g(x) - M|}{|g(x)||M|} < \frac{\epsilon}{2}$

so we want

$$|f(x) - L| < \frac{|g(x)|}{2} \epsilon$$

and we want

$$|g(x) - M| < \frac{|g(x)||M|}{2|L|} \epsilon$$

98

For the first condition, we need a lower bound on $|g(x)|$. We get this by choosing δ_1 so that if $0 < |x - a| < \delta_1$ we have $|g(x) - M| < \frac{|M|}{2}$, which ensures that $|g(x)| \geq \frac{|M|}{2}$. $[|g(x) - M| + |-g(x)| \geq |-M| = |M|$ by the triangle inequality, so $|g(x)| = |-g(x)| \geq |M| - |g(x) - M| \geq |M| - \frac{|M|}{2} = \frac{|M|}{2}$ – this is the calculation using the Triangle Inequality that I couldn't come up with live in class – it might be possible to make it shorter and sweeter]

Choose δ_2 such that if $0 < |x - a| < \delta_2$ then $|f(x) - L| < \frac{|M|}{4}\epsilon$

For the second condition, we actually need to allow for the possibility that $L = 0$, so we first replace $|L|$ in the denominator with $|L| + 1$. We also need a lower bound on $|g(x)|$, which we already have.

Choose δ_3 such that if $0 < |x - a| < \delta_3$ then $|g(x) - M| < \frac{|M|^2}{4(|L|+1)}\epsilon$.

Proof resumes: Choose $\delta_1, \delta_2, \delta_3$ as described in the scratch work and let δ be the smallest of these three numbers.

Choose x arbitrarily. Assume $0 < |x - a| < \delta$. It follows in particular that $0 < |x - a| < \delta_1$, so $|g(x)| < \frac{|M|}{2}$ by the argument given in the scratch work.

Now
$$|\frac{f(x)}{g(x)} - \frac{L}{M}| \leq \frac{|f(x) - L|}{g(x)} + \frac{|L||g(x) - M|}{|g(x)||M|}$$

by calculations in the scratch work

and this is less than

$$\frac{2(|f(x) - L|)}{|M|} + \frac{2|L||g(x) - M|}{|M||M|}$$

by the lower bound on $|g(x)|$ given

which is less than

$$\frac{2(\frac{|M|}{4}\epsilon)}{|M|} + \frac{2|L|\frac{|M|^2}{4(|L|+1)}\epsilon}{|M||M|}$$

which simplifies to $\frac{\epsilon}{2} + \frac{|L|}{|L|+1}\frac{\epsilon}{2} < \epsilon$

6.6.2 Nonexistence of Limits – A Lemma and Examples

The statement $\lim_{x \to a} \neq L$ converts by logical transformations that you should be able to think through to the following form:

For some $\epsilon > 0$, for all $\delta > 0$, for some x, $0 < |x-a| < \delta$ and $|f(x)-L| \geq \epsilon$

You should be able to figure out why this works, and I would regard it as fair to ask you on an exam to write out the form of $\lim_{x \to a} \neq L$ without any explicit negations, which I have just given: you should be able to do this by understanding how to negate quantified statements, how to negate an implication, and how to negate an order statement.

This is not really what we want for proving nonexistence of limits, though. The problem with it is the mention of a specific L. We don't want to show that a limit expression isn't equal to a specific L in such a proof, but that it is meaningless (is not equal to *any L*).

For this purpose, it is useful to prove the following

Lemma: If $\lim_{x \to a} f(x) = L$ then for any $\epsilon > 0$, there is a $\delta > 0$ such that for all x, y, if $0 < |x - a|, |y - a| < \delta$ then $|f(x) - f(y)| < \epsilon$.

Proof of Lemma: Let $\epsilon > 0$ be chosen arbitrarily.

Let δ be chosen so that if $0 < |z - a| < \delta$, for any z, then $|f(z) - L| < \frac{\epsilon}{2}$. We can do this by the assumed limit statement.

Let x, y be chosen so that $0 < |x - a|, |y - a| < \delta$.

Then

$$|f(x) - f(y)| = |(f(x) - L + L - f(y)| \leq |f(x) - L| + |f(y) - L|$$

[Triangle inequality]
$$< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon,$$

which is what we need.

Contrapositive form of the Lemma: If it is not true that there is a $\delta > 0$ such that for all x, y, if $0 < |x - a|, |y - a| < \delta$ then $|f(x) - f(y)| < \epsilon$ then $\lim_{x \to a} f(x) \neq L$ for any L at all – notice that the hypothesis of this implication does not mention L. Applying logical transformations, we get this:

Nonexistence of Limits Lemma: If for some $\epsilon > 0$, for all $\delta > 0$, there are numbers x, y such that $0 < |x-a|, |y-a| < \delta$ and $|f(x)-f(y)| \geq \epsilon$, then $\lim_{x \to a} f(x)$ does not exist (because by the contrapositive form of the Lemma above it cannot be equal to any L).

Example 1: Show that $\lim_{x \to 0} \frac{|x|}{x}$ does not exist.

To be added.

Example 2: Show that $\lim_{x \to 0} \sin(\frac{1}{x})$ does not exist.

To be added.

Theorem: If $\lim_{x \to a} f(x) = L \neq 0$ and $\lim_{x \to a} g(x) = 0$ then $\lim_{x \to a} \frac{f(x)}{g(x)}$ does not exist.

To be added.

6.6.3 Homework assigned 10/30/2014 – boo!

This assignment is due on Friday, November 7. Start on it right away; it is very probable that another problem set will be posted before the 7th.

From Spivak chapter 5, do problem 8, 9, 10, 12a, 17b, 18, 19, 38

6.7 Spivak Chapter 6 – Continuity

You should make sure you know how to write out the proof that (for example) $f + g$ is continuous if f and g are continuous (at a particular a).

Theorem: If f is continuous at a and g is continuous at a then $f + g$ is continuous at a.

Proof: Suppose that f is continuous at a (this means $\lim_{x \to a} f(x) = f(a)$) and g is continuous at a (this means $\lim_{x \to a} g(x) = f(a)$)

Our goal is to prove that $f + g$ is continuous at a: this means that we want to prove $\lim_{x \to a}(f + g)(x) = (f + g)(a)$

$\lim_{x \to a}(f + g)(x) =$ [definition of $f + g$]
$\lim_{x \to a}[f(x)+g(x)] =$[addition property of limits, if the two limits exist, which we will see that they do]
$\lim_{x \to a} f(x) + \lim_{x \to a} g(x) =$[assumptions]
$f(a) + g(a) =$[definition of $f + g$]
$(f + g)(a)$.

Of course there are similar proofs for the other operations.

The next theorem that Spivak gives is the important theorem that the composition of two continuous functions is continuous. I prove a related result; you are assigned the proof of a different related result in the chapter 6 homework.

The result that I prove has to do with changes of variable in limits. A typical application which appears in Math 170 (without any real formal justification) is something like this: $\lim_{x \to 0} \frac{\sin(3x)}{3x} = 1$ by a change of variables: set $u = 3x$ and notice that $u \to 0$ as $x \to 0$, so we can write this $\lim_{u \to 0} \frac{\sin(u)}{u} = 1$ and this fact is known.

This kind of reasoning is generally not supported in the Math 170 book that uses it at all. I state a theorem. The theorem I stated and proved in class was *incorrect* (slightly).

Theorem (change of variables in limits): Suppose that $\lim_{x \to a} g(x) = b$, $\lim_{x \to b} f(x) = L$, and either $f(b) = L$ or there is a δ such that $g(x) \neq b$ for any x with $0 < |x - a| < \delta$ (this is the needed correction; I had g continuous, which does not help with the actual problem). Then $\lim_{x \to a} f(g(x)) = L$.

Discussion: From this theorem we can justify this kind of reasoning: suppose $u = g(x)$ and as $x \to a$, we have $u \to b$, and we want to compute $\lim_{x \to a} f(g(x))$: we change variables to conclude that $\lim_{u \to b} f(u)$ is an equivalent limit expression. If f is continuous at b this always works: otherwise we have to be sure that u does not take on the value b for x too close to a (except exactly at a, where we do not care what happens).

Proof: We choose $\epsilon_0 > 0$ arbitrarily. Because $\lim_{x \to b} f(x) = L$ we know that there is δ_1 such for any x, if $0 < |x - b| < \delta_1$, we will have $|f(x) - L| < \epsilon_0$. We call this (*).

Because $\lim_{x \to b} g(x) = b$, we know that we can choose δ_2 such that for any x if $0 < |x - a| < \delta_0$, we will have $|g(x) - b| < \delta_1$ (notice that here we are using a number introduced as a δ as the ϵ value for another limit fact).

If we have a δ_3 such that for all x with $0 < |x - a| < \delta_3$, $g(x) \neq b$, we define δ_0 as the minimum of δ_2 and δ_3; otherwise we define δ_0 as δ_2. Note that this implies that if $0 < |x - a| < \delta_0$ and $g(x) = b$, we must have $f(b) = L$.

So if $0 < |x - a| < \delta_0$ we have $|g(x) - b| < \delta_1$. If $g(x) \neq b$, we have $0 < |g(x) - b| < \delta_1$, so we have $|f(g(x)) - L| < \epsilon_0$: this results from replacing x with $g(x)$ in (*). As noted above, if we have $g(x) = b$ we must have $f(b) = L$ so $|f(g(x)) - L| = 0 < \epsilon_0$.

It is a corollary of this theorem that if f is continuous at $g(a)$ and g is continuous at a, then $f \circ g$ is continuous at a.

A theorem which will have important uses in chapter 8, and which you should know how to prove if called upon to do so on the coming test, is the following:

Theorem: If f is continuous at a and $f(a) > 0$, then there is a δ such that for all x in $(a - \delta, a + \delta)$ we have $f(x) > 0$.

Proof: Suppose that f is continuous at a and $f(a) > 0$.

By definition of continuity, $\lim_{x \to a} f(x) = f(a)$.

This means that there is a δ such that for any x if $0 < |x-a| < \delta$ we have $|f(x) - f(a)| < \frac{f(a)}{2}$, from whice we get $f(a) - \frac{f(a)}{2} < f(x) < f(a) + \frac{f(a)}{2}$, from which we see that $0 < \frac{f(a)}{2} < f(x)$:

Suppose that x is in $(a - \delta, a + \delta)$. Either $x = a$, in which case $f(x) = f(a) > 0$, or $0 < |x - a| < \delta$, in which case we have just shown $0 < \frac{f(a)}{2} < f(x)$. In either case, $f(x) > 0$, so we have proved the theorem.

Finally, I write up the evil example of a function which is continuous at only one point. The issue here is to be aware that the concept which corresponds to "being able to draw the graph without picking up the pencil from the paper" (the intuitive notion of continuity) corresponds to continuity at every point in an interval; continuity at a single point can occur in situations where no connected curve is present in the graph at all, as in this example.

Consider the function f defined by "f(x)=0 if x is irrational and x if x is rational".

First Claim: this function is continuous at 0.

What we need to prove is $\lim_{x \to 0} f(x) = 0$, that is, for each $\epsilon > 0$ there is a $\delta > 0$ such that if $0 < |x| < \delta$, then $|f(x)| < \epsilon$.

Choose an arbitrary $\epsilon_0 > 0$. Let $\delta_0 = \epsilon_0$.

Choose an arbitrary x, and suppose $0 < |x| < \delta_0$ (*). Our aim is to show $|f(x)| < \epsilon_0$.

If x is rational, then $f(x) = x$ and we have $|f(x)| = |x| < \delta_0 [\text{by }(*)] = \epsilon_0$.

If x is irrational then $f(x) = 0$ and we have $|f(x)| = 0 < \delta_0 = \epsilon_0$ just because ϵ_0 is positive.

This completes the proof that f is continuous at 0.

Second Claim: f is not continuous anywhere else.

Suppose for the sake of a contradiction that f is continuous at some $a \neq 0$.

Then we have $\lim_{x \to a} f(x) = L$ for some L.

It follows that there is a δ_1 such that for any x if $0 < |x - a| < \delta_1$, it follows that $|f(x) - L| < \frac{|a|}{3}$.

Define δ_0 as $\min(\delta_1, \frac{|a|}{3})$.

Now if x is rational and $0 < |x - a| < \delta_0$, it follows that $|f(x) - L| = |x - L| < \frac{|a|}{3}$, so further we have $|a - L| \leq |a - x| + |x - L| = |x - a| + |x - L| < \frac{2|a|}{3}$ ($|x - a|$ is bounded because of the additional bound we put on δ_0).

If x is irrational and $0 < |x - a| < \delta_0$, it follows that $|f(x) - L| = |L| < \frac{|a|}{3}$.

Now $|a| \leq |L| + |a - L| < \frac{|a|}{3} + \frac{2|a|}{3} = |a|$, which is absurd.

A discussion of just how I came up with $\frac{|a|}{3}$ as the bad ϵ would be useful.

6.8 Discussion of the Intermediate and Extreme Value Theorems

Add some stuff here...

6.9 Least Upper Bounds and All That

In this section we introduce the final property P13 which completes the statement of our theory of the real numbers.

Definition: Let A be a set of numbers.

> We say that A is *bounded above* iff there is a number M such that $(\forall a \in A.a \leq M)$. Such a number M is called an *upper bound* for A. Notice that "A is bounded above" means exactly the same thing as "there is an upper bound for A" or "A has an upper bound".

> We say that A is *bounded below* iff there is a number m such that $(\forall a \in A.a \geq m)$. Such a number m is called a *lower bound* for A. Notice that "A is bounded below" means exactly the same thing as "there is a lower bound for A" or "A has a lower bound".

Definition: Let A be a set of numbers which is bounded above. We say that a number B is a *least upper bound* of A iff B is an upper bound for A and for any upper bound M of A, $B \leq M$.

> Let A be a set of numbers which is bounded below. We say that a number b is a *greatest lower bound* of A iff b is a lower bound of A and for any lower bound m of A, $b \geq m$.

It should be clear that a set can have at most one least upper bound and at most one greatest lower bound. This justifies the introduction of notations $\sup(A)$ or $\mathtt{lub}(A)$ for the least upper bound of A and $\inf(A)$ or $\mathtt{glb}(A)$ for the greatest lower bound of A.

Observation: Notice that if a set A has a largest element $\max(A)$, then $\max(A)$ is an upper bound for A (note the use of $\leq$ in the definition) and obviously is the least upper bound. Similarly, if A has a smallest element $\min(A)$ then $\min(A)$ is a lower bound for A, and obviously the greatest lower bound.

It is important to reiterate a point. You all know this but many of you will forget it in your anxieties when faced with a proof to write. A set of numbers which is bounded above does not necessarily have a maximum element. For example, consider the set of negative numbers, or any open interval such as $(0, 1)$. A set of numbers which is bounded below does not necessarily have a minimum element. Consider the set of positive numbers, or $(0, 1)$ again, or $\{\frac{1}{n} \mid n \in \mathbb{Z}^+\}$.

But the system of real numbers has a property which is almost as good.

Axiom (P13): Any set of real numbers which is nonempty and bounded above has a least upper bound.

Theorem (P13b): STAR Any set of real numbers which is nonempty and bounded below has a greatest lower bound.

Proof: Let A be a nonempty set which is bounded below. Define $-A$ as the set
$\{x \mid -x \in A\}$, or equivalently $\{-x \mid x \in A\}$. Obviously $-A$ is nonempty.

We claim first that $-A$ is bounded above. Let m be a lower bound for A. We claim that $-m$ is an upper bound for $-A$: let $x \in -A$, and we need to show $x \leq -m$. This is equivalent to $-x \geq m$, which is true because $-x \in A$, and m is a lower bound for A.

Thus by P13 $-A$ has a least upper bound B.

We claim that $-B$ is the greatest lower bound of A. There are a couple of things to show to verify this. First, we need to verify that $-B$ is a lower bound of A. Let $x \in A$: we need to show $x \geq -B$;

this is equivalent to $-x \leq B$, which is true because $-x \in -A$ and B is an upper bound for $-A$. Second, we need to verify that for any lower bound m of A, $-B \geq m$. We have already shown that for any lower bound m of A, $-m$ is an upper bound of A. We know then that $B \leq -m$ because B is the least upper bound of $-A$; this is equivalent to $-B \geq m$ which is what we needed to show.

Observation: P13 is *not* true in the system of rational numbers; this shows that it does not follow from P1-12. For example the set $\{x \in \mathbf{Q} \mid x^2 < 2\}$ has an upper bound in the system of rational numbers (2 would work) but it has no *rational* least upper bound. A rational number q will be an upper bound for this set iff $q > 0$ and $q^2 > 2$. Whatever q you pick, you can find a smaller q' with the same property (something which is rational and whose square is even closer to 2). Proving this is a tricky bit of algebra, but it's not at all hard to believe. Of course in the full system of real numbers this set is bounded above and has the least upper bound $\sqrt{2}$ (which is not rational).

6.10   Notes from November 5 2014

Discussion from fall 2014: I did a full proof of the observation above.

Theorem: P13 is false in the rational numbers.

Proof: Let A be the set of all rational numbers q such that $q^2 < 2$. We claim that this set has no least upper bound in the rationals.

Suppose for the sake of a contradiction that M was a rational least upper bound for the set A.

We know that $M^2 \neq 2$ (we have proved already that there is no rational square root of 2).

So either $M^2 > 2$ or $M^2 < 2$.

We know that the function $f(x) = x^2$ is continuous so $\lim_{x \to M} x^2 = M^2$. We can find a δ such that if $0 < |x - M| < \delta$ we have $|x^2 - M^2| < \frac{|M^2 - 2|}{2}$, by the definition of limit. Choose such a δ.

If $M^2 > 2$, let $x = M - \frac{\delta}{2}$. Since $|x^2 - M^2| < \frac{|M^2 - 2|}{2}$, we have $x^2 > 2$ as well. Since $x^2 > 2 > a^2$ for any $a \in A$, and certainly x is positive, we have $x > a$ for all $a \in A$ – but then x is an upper

bound for A, and less than M which is supposed to be the least upper bound for A, which is a contradiction.

If $M^2 < 2$, let $x = M + \frac{\delta}{2}$. Since $|x^2 - M^2| < \frac{|M^2 - 2|}{2}$, we have $x^2 < 2$ as well. But then $x \in A$ and $x > M$, so M is not an upper bound for A, which is a contradiction.

I at least suggested the reasons for equivalence of P13 with a statement about intervals.

Definition: An interval is a set defined in one of the following ways (where $a \leq b$ are numbers):

1. $[a, b] = \{x : a \leq x \leq b\}$
2. $(a, b] = \{x : a < x \leq b\}$
3. $[a, b) = \{x : a \leq x < b\}$
4. $[a, b] = \{x : a < x < b\}$
5. $[a, \infty) = \{x : a \leq x\}$
6. $(a, \infty) = \{x : a < x\}$
7. $[-\infty, b) = \{x : x \leq b\}$
8. $(-\infty, b) = \{x : x < b\}$
9. $(-\infty, \infty) = \mathbb{R}$

Definition: A set C of numbers is *convex* iff for every $x, y \in C$, for any z, if $x < z < y$ then $z \in C$. Equivalently, for each pair of numbers x, y such that $x < y$ and both x and y belong to C, $[x, y]$ is a subset of C.

Theorem: In the presence of axioms P1-12, axiom P13 is equivalent to the statement that every convex set is an interval. [Stewart's calculus book, at least in some editions, states the Least Upper Bound Property by accident, because it includes the statement that the convex sets are exactly the intervals; most calculus books do not officially state this property, or at least if they do, they do not state it in their preliminaries but in an appendix. I enjoyed finding this!]

Proof: To be added

Here is an important lemma which very likely already appears in the notes below already, but I'll make sure it is present by putting it here.

Lemma: Let A be a nonempty set of reals which is bounded above. Let M be the least upper bound of A. It is possible for M to belong to A: in this case, M is the largest element of A. It is possible for M not to belong to A: in this case, for every $\epsilon > 0$, there is $a \in A$ such that $M - \epsilon < a < M$.

Proof of Lemma: The interval [0,1] is an example of a set whose least upper bound 1 belongs to it. If A is a nonempty set bounded above and M is its least upper bound and $M \in A$, there cannot be $a \in A$ such that $a > M$ because M is an upper bound for A, so we have $M \in A$ and $a \leq M$ for any $a \in A$ by trichotomy, and this is what it means for M to be the largest element of A.

The interval (0,1) is an example of a set whose least upper bound 1 does not belong to it. Suppose that A is a nonempty set bounded above and $M \notin A$ is its least upper bound. Let $\epsilon > 0$ be chosen arbitrarily. Since $M - \epsilon < M$, $M - \epsilon$ cannot be an upper bound for A, so we can choose $a \in A$ such that $M - \epsilon < a$. We also have $a \leq M$ because M is an upper bound for A, but actually $a < M$ because $M \notin A$ is distinct from $a \in A$. So we have shown that there is $a \in A$ such that $M - \epsilon < a < M$, which is what we needed to show.

Observation: Notice that if A is a nonempty set bounded above and $M \notin A$ is the least upper bound for A, we can choose $a_1 \in A$ such that $M - 1 < a_1 < M$, then choose $a_2 \in A$ such that $a_1 < a_2 < M$ (by the previous lemma with $\epsilon = |M - a_1|$) and continue indefinitely, choosing $a_{i+1} \in A$ such that $a_i < a_{i+1} < M$ for each natural number i. So we can find infinitely many points of A as close to M as we want.

Same thing for greatest lower bounds: If M is the greatest lower bound of a nonempty set A, then either M is the least element of A or $M \notin A$ and for every $\epsilon > 0$ there is $a \in A$ with $M < a < M + \epsilon$. Both situations are possible. The proof is easy to write following the one above.

Here is a very useful theorem and corollary. Though you might think it is, this is NOT a consequence of P1-12!

Theorem: For each number x, there is a natural number n such that $x < n$.

Proof: Suppose otherwise. That is, suppose that there is a fixed number x such that for every natural number n, $x < n$. If there were such a number, then the set $\mathbb{N}$ of natural numbers would be bounded above, and since it is certainly nonempty, it would have a least upper bound M.

By the Lemma above, there would be a natural number N such that $M - 1 < N < M$. Then observe that $M < N + 1 < M + 1$ by properties of order, $N + 1$ is a natural number because natural numbers are closed under successor, and this contradicts the assumption that M is an upper bound for the natural numbers, as $M < N + 1$.

We make an observation about the natural numbers. The only properties of the set of natural numbers which we used above are that this set is nonempty and that it is closed under successor. In fact, we have not defined the set of natural numbers at all! We give the definition for completeness – it is rather tricky.

Definition: A set of numbers I such that $1 \in I$ (or such that $0 \in I$ if we want 0 be be a natural number) and $(\forall x \in I.x + 1 \in I)$ [I is closed under successor] is called an *inductive set*.

A number n is a natural number iff it belongs to all inductive sets (notice that the choice we make in the definition of inductive sets determines whether 0 is a natural number or not).

The collection of all natural numbers is called $\mathbb{N}$.

If you look carefully, you will notice that we have made mathematical induction true by definition of $\mathbb{N}$!

6.11 Proofs of the Major Theorems

Theorem: Let f be a function which is continuous on an interval $[a, b]$ and let $f(a) < 0$ and $f(b) > 0$. Then there is c in (a, b) such that $f(c) = 0$. (this is chapter 7 theorem 1, from which we have already seen how to deduce the full Intermediate Value Theorem).

Proof: We are going to prove this using the Least Upper Bound Property P13. The magic lies in the choice of the appropriate set... Let S be

defined as $\{x \in [a, b] \mid (\forall y \in [a, x].f(y) < 0)\}$. (This is not exactly the same as the definition $S = \{x \in [a, b] \mid (\forall y \in [a, x).f(y) < 0)\}$I used in class; the proof is very similar to what I did in class, but I think slightly easier). S is the set of all numbers at which the value of the function is negative, and for which all values of the function at smaller numbers in $[a, b]$ is also negative. Notice (this makes our life easier) that it is quite clear that if $x \in S$ and $a \le y < x$, then y is also in S: if all values of the function at numbers less than x but $\ge a$ are negative, then clearly the same will be true for any y between a and x.

The set S is nonempty because obviously $a \in S$ (because $f(a) < 0$). The set S is bounded above by b. So by P13 S has a least upper bound which we will call c. We claim that $c \in (a, b)$ and $f(c) = 0$ (these are goal statements and require proof).

To show that $c \in (a, b)$ we need to show that $c \ne a$ and $c \ne b$. To show that $f(c) = 0$ we show that $\neg f(c) > 0$ and $\neg f(c) < 0$. The rest of the proof is contained in the proofs of these four claims.

Claim 1: $c \ne a$

Because $\lim_{x \to a+} f(x) = f(a)$, there is a $\delta > 0$ such that for any x with $a < x < a + \delta$ we have $|f(x) - f(a)| < \frac{|f(a)|}{2}$, from which it follows that $f(a) - \frac{|f(a)|}{2} < f(x) < f(a) + \frac{|f(a)|}{2} = f(a) - \frac{f(a)}{2} = \frac{f(a)}{2} < 0$. Remember that $f(a) < 0$, so $|f(a)| = -f(a)$. This implies that any $x < a + \delta$ belongs to S, because this shows that at x and everywhere below x the function f takes on negative values. But this means that for example $a + \frac{\delta}{2} \in S$, so a is not an upper bound for S, so $a \ne c$.

Claim 2: $c \ne b$

Because $\lim_{x \to b-} f(x) = f(b)$, there is a $\delta > 0$ such that for any x with $b - \delta < x < b$ we have $|f(x) - f(b)| < \frac{|f(b)|}{2}$, from which it follows that $0 < \frac{f(b)}{2} = f(b) - \frac{|f(b)|}{2} < f(x) < f(b) + \frac{|f(b)|}{2}$. Remember that $f(b) > 0$, so $|f(b)| = f(b)$. This implies that no $x > b - \delta$ belongs to S, because this shows that at x the function f takes on a positive value. This means not only that x is not in S but that x is an upper bound for S (if x were less than any element of S it would also be in S as we showed above). This

means that for example $b - \frac{\delta}{2}$ is an upper bound for S, and this is less than b so b cannot be the least upper bound for S.

Claim 3: $\neg f(c) > 0$

Suppose $f(c) > 0$ for the sake of a contradiction. We know that $c \in (a, b)$ since we established the two claims above, so we know that $\lim_{x \to c} f(x) = f(c)$, so we know that there is $\delta > 0$ such that for all x such that $0 < |x - c| < \delta$ we have $|f(x) - f(c)| < \frac{|f(c)|}{2} = \frac{f(c)}{2}$. In fact this is true for all x in $(c - \delta, c + \delta)$ because it is true for $x = c$ as well. Now this implies that $0 < \frac{f(c)}{2} = f(c) - \frac{|f(c)|}{2} < f(x) < f(c) + \frac{|f(c)|}{2}$. This means that the value of f is positive everywhere in $(c - \delta, c + \delta)$. This means that $c - \frac{\delta}{2}$ is not in S, so is an upper bound for S, which implies that c cannot be the least upper bound for S, since we just found a smaller upper bound for S.

Claim 4: $\neg f(c) < 0$

Suppose $f(c) < 0$ for the sake of a contradiction. We know that $c \in (a, b)$ since we established the two claims above, so we know that $\lim_{x \to c} f(x) = f(c)$, so we know that there is $\delta > 0$ such that for all x such that $0 < |x - c| < \delta$ we have $|f(x) - f(c)| < \frac{|f(c)|}{2} = -\frac{f(c)}{2}$. In fact this is true for all x in $(c - \delta, c + \delta)$ because it is true for $x = c$ as well. Now this implies that $f(c) - \frac{|f(c)|}{2} < f(x) < f(c) + \frac{|f(c)|}{2} = f(c) - \frac{f(c)}{2} = \frac{f(c)}{2} < 0$. This means that the value of f is negative everywhere in $(c - \delta, c + \delta)$. We claim further than the value of f is negative at any x which is less than c (and greater than or equal to a): any $x < c$ is not an upper bound for S, so it is less than some $y \in S$, so it is in S itself, so the value of x there is negative. But now we can see that $c + \frac{\delta}{2} \in S$: we have shown that the value of f at $c + \frac{\delta}{2}$ and everywhere below $c + \frac{\delta}{2}$ in the interval $[a, b]$ is negative. And this is absurd, because c is supposed to be an upper bound for S.

Theorem: If f is continuous on $[a, b]$, then there is an M such that for all $x \in [a, b]$, $f(x) \leq M$: that is, the set of values of f on $[a, b]$ is bounded above.

Proof: We will prove this using P13, so we will define a suitable set which has a least upper bound...

The set S is defined as $\{x \in [a, b] \mid (\exists M.(\forall y \in [a, x].f(y) \leq M))\}$. In other words, x belongs to S iff x is in $[a, b]$ and the set of values of f on the interval $[a, x]$ is bounded.

The set S is nonempty because $a \in S$: the set of all values of f on $[a, a]$ is just $\{f(a)\}$, which has $f(a)$ itself as an upper bound.

The set S is bounded above because all elements of S are in $[a, b]$, so b is an upper bound.

Thus by P13 S has a least upper bound.

The set S is downward closed: that is, if $x \in S$ and $a < y < x$ then $y \in S$: if $x \in S$, this means there is an upper bound on values of f on the interval $[a, x]$, which will also be an upper bound for values of f on the smaller interval $[a, y]$, so $y \in S$ as well. This means that for any z which is in $[a, b]$ and which is not an upper bound for S we can deduce $z \in S$: because z is not an upper bound for S, there is $w > z$ which belongs to S, so $z \in S$ because S is downward closed.

We prove each of a sequence of claims.

Claim 1 $\sup(S) \neq a$: We know that $\lim_{x \to a+} f(x) = f(a)$ (part of the assertion that f is continuous on $[a, b]$).

Fix an $\epsilon > 0$. Choose a $\delta > 0$ such that for each x such that $a < x < a + \delta$, $|f(x) - f(a)| < \epsilon$. It follows that for any x with $a < x < a + \delta$, $f(x) < f(a) + \epsilon$. Further, it follows that for any x in $[a, a + \frac{\delta}{2}]$, $f(x) < f(a) + \epsilon$. Thus the set of values of f on $[a, a + \frac{\delta}{2}]$ is bounded above, so $a + \frac{\delta}{2} \in S$, so a is not an upper bound for S, so $a \neq \sup(S)$.

Claim 2 $\sup(S) \notin (a, b)$: Suppose otherwise for the sake of a contradiction, that is, suppose that $\sup(S) = c$, where $a < c < b$.

By continuity assumptions about f, $\lim_{x \to c} f(x) = f(c)$.

Fix $\epsilon > 0$. Choose $\delta > 0$ such that for any x with $0 < |x - a| < \delta$ we have $|f(x) - f(c)| < \epsilon$. In fact, we have for any x with $c - \delta < x < c + \delta$ that $|f(x) - f(c)| < \epsilon$ (because this is clearly true for c as well). It follows that for any x in $(c - \delta, c + \delta)$, $f(x) < f(c) + \epsilon$. Because $c - \delta < c$, the least upper bound of S, this means that $c - \delta$ is not an upper bound for S, so $c - \delta \in S$ by the discussion above. This means that there is an upper bound M for the set

113

of values of f on $[a, c - \delta]$. We know that $f(c) + \epsilon$ is an upper bound for the set of values of f on $(c - \delta, c + \delta)$ and so also on $(c - \delta, c + \frac{\delta}{2}]$. Thus $\max(M, f(c) + \epsilon)$ is an upper bound for values of f on $[a, c + \frac{\delta}{2}]$, so $c + \frac{\delta}{2} \in S$, which is a contradiction because c is supposed to be an upper bound for S.

Claim 3 $\sup(S) = b$**:** The Claim requires no proof but it does require a remark to indicate why we are not finished. b is an upper bound for S and we have seen that nothing smaller can be, so b is the least upper bound for S. What we need to be done is to actually show that $b \in S$.

Claim 4: $b \in S$ **(the theorem):** We know that $\lim_{x \to b-} f(x) = f(b)$ by continuity assumptions about f. Fix $\epsilon > 0$. Choose $\delta > 0$ such that for any x with $b - \delta < x < b$ we have $|f(x) - f(b)| < \epsilon$, so $f(x) < f(b) + \epsilon$. Now $b - \delta \in S$ because $b - \delta$ is not an upper bound for S. Thus there is an M which is an upper bound for values of f on $[a, b - \epsilon]$. This means that $\max(M, f(b) + \epsilon)$ is an upper bound for values of f on $[a, b - \epsilon]$, on $(b - \epsilon, b)$, and (obviously) at b, so on all of $[a, b]$, which proves the theorem.

Theorem: If f is continuous on $[a, b]$, then there is $c \in [a, b]$ such that for all $x \in [a, b]$, $f(x) \le f(c)$.

Proof: Let $S = \{f(x) \mid x \in [a, b]\}$, the set of all values of the function f on the interval $[a, b]$. S is nonempty (it contains $f(a)$ for example) and bounded above by the previous theorem, so S has a least upper bound which we will call B. If B were an element of S it would be the largest element of S, that is, an $f(c)$ greater than or equal to all values $f(x)$, and we have assumed that there is no such $f(c)$.

Suppose that $B \notin S$. This means that $f(x) < B$ for all $x \in [a, b]$. Thus $g(x) = \frac{1}{B - f(x)}$ defines a function which is defined and continuous on the interval $[a, b]$. We claim that the values of g on $[a, b]$ are not bounded above. Suppose $g(x) \le C$ for all $x \in [a, b]$. It would follow that $\frac{1}{B - f(x)} \le C$, so $B - f(x) \ge \frac{1}{C}$, so $B - \frac{1}{C} \le f(x)$ for all $x \in [a, b]$. But this is impossible: $B - \frac{1}{C}$ would then be an upper bound for the set S of all values of f on $[a, b]$, and B is the least upper bound of this set. So the alternative $B \in S$ must hold, so $B = f(c)$ for some $c \in [a, b]$, and $B = f(c) \ge f(x)$ for all

$x \in [a, b]$.

6.12 Chapter 8 Homework Assignment – assigned 11/12/2014

From Spivak do problems 1, 2 (I outlined this in class), 5, 6, 7, 8, 12, 14

Some of these are **brutally hard**. You might think they **all** are. You should come talk to me about them when you have discovered this.

This assignment is due on the Friday before break.

6.13 Sequences

Definition: A sequence is a function with domain $\mathbb{Z}^+$ (the positive integers). 0 is often included in the domain. If a is a sequence, it is usual to write a_n instead of $a(n)$.

There is a notion of limit proper to sequences.

Definition: We say that $\lim_{n\to\infty} a_n = L$ just in case for each $\epsilon > 0$ there is an N such that for all $n > N$, $|a_n - L| < \epsilon$. A sequence $\{a_n\}$ is said to converge if $\lim_{n\to\infty} a_n = L$ for some L (if its limit exists).

Notice that we also have limits at infinity for functions of the reals: $\lim_{x\to\infty} f(x) = L$ just in case for each $\epsilon > 0$ there is an N such that for all $n > N$, $|f(x) - L| < \epsilon$.

These notions are not quite the same. It is true that if $a_n = f(n)$ and $\lim_{x\to\infty} f(x) = L$, then $\lim_{n\to\infty} a_n = L$. The converse is not true: the sequence whose nth term is $\sin(n\pi)$ has all terms 0 so certainly has limit 0. but $\lim_{x\to\infty} \sin(\pi x)$ is undefined (the function oscillates between -1 and 1 but happens to hit 0 at every whole number).

Theorem: STAR The set of natural numbers is unbounded.

Proof: It may seem very surprising that we have to prove this at all. It's beyond the scope of this course, but it is quite possible to define a structure satisfying axioms P1-P12 in which there are "numbers" greater than all natural numbers. But P13 allows us to avoid this weird situation. Note that this result is the entry point for chapter 8 problem 5.

The set of natural numbers $\mathbb{N}$ is defined as follows: a set I is *inductive* iff $0 \in I$ and $(\forall x. x \in I \rightarrow x + 1 \in I)$: a number n is a natural number just in case for every inductive set I, $n \in I$. It is obvious how to modify the definition if we want 0 to be a natural number. The interesting thing here is that we actually can't define $\mathbb{N}$ without referring to sets, and also that the natural numbers are defined precisely so as to make math induction work.

Suppose for the sake of a contradiction that $\mathbb{N}$ is bounded above. Clearly $\mathbb{N}$ is nonempty (1 is an element) so if it is bounded above it has a least upper bound B. Now observe that $B - 1$ is not an upper bound for $\mathbb{N}$ (because B is the least upper bound) so there is a natural number N with $B - 1 < N \leq B$. But clearly then $N + 1$ is a natural number with $B < N + 1 \leq B + 1$ (adding 1 to both sides of each of the previous inequalities) so B fails to be an upper bound for $\mathbb{N}$, which is a contradiction.

Theorem: STAR $\lim_{x \to \infty} \frac{1}{n} = 0$

Proof: Fix $\epsilon > 0$. We want to find N such that if $n > N$ we will have $|\frac{1}{n} - 0| < \epsilon$. Choose a natural number $N > \frac{1}{\epsilon}$ (you can do this by the previous theorem; this is why we had to prove that theorem). Then $n > N$ implies $\frac{1}{n} < \frac{1}{N} < \epsilon$ and since $n > 0$ we have $|\frac{1}{n} - 0| = \frac{1}{n} < \frac{1}{N} < \epsilon$. This is what we need.

The basic limit laws are essentially the same for limits of sequences. We prove one of them: notice that we want maxima of N's where we want minima of δ's in proofs about limits of functions.

Theorem: STAR If $\lim_{n \to \infty} a_n = L$ and $\lim_{n \to \infty} b_n = M$ then $\lim_{n \to \infty} a_n + b_n = L + M$.

Proof: This proof will look almost entirely familiar except for the mentioned detail about N's.

Assume $\lim_{n \to \infty} a_n = L$ and $\lim_{n \to \infty} b_n = M$ Fix $\epsilon_0 > 0$. By the assumptions there is N_1 such that for all $n > N_1$ we have $|a_n - L| < \frac{\epsilon_0}{2}$, and N_1 such that for all $n > N_2$ we have $|b_n - M| < \frac{\epsilon_0}{2}$. Set $N_3 = \max(N_1, N_2)$. The point here is that if $n > N_3$ then $n > N_1$ *and* $n > N_2$ so both of the closeness conditions we want will hold.

Now choose n arbitrarily. Assume $n > N_3$. It follows that $n > N_1$ so $|a_n - L| < \frac{\epsilon_0}{2}$ and $n > N_2$ so $|b_n - M| < \frac{\epsilon_0}{2}$, so $|(a_n + b_n) - (L + M)| = |(a_n - L) + (b_n - M)| \le (\Delta)|a_n - L| + |b_n - M| < \frac{\epsilon_0}{2} + \frac{\epsilon_0}{2} = \epsilon_0$.

Here is a theorem we quote but do not prove in Math 175, which is another statement which is actually equuivalent to P13.

Definition: A sequence is nondecreasing iff $a_n \le a_{n+1}$ for each n. Notice that this implies that $a_m \le a_n$ if $m < n$. If we swap out $\le$ for $<, \ge, >$ respectively we get the notions of (strictly) increasing, nonincreasing, and (strictly) decreasing sequences. A sequence which is either nonincreasing or nondecreasing is called a monotone sequence.

Definition: We say that a sequence is bounded above (below) if the set $\{a_n \mid n \in \mathbb{N}\}$ of its values is bounded above (below). A sequence is said to be bounded if it is bounded both above and below.

Theorem: STAR A nondecreasing sequence which is bounded above converges.

Proof: Assume that $\{a_n\}$ is a nondecreasing sequence which is bounded above. Let $S = \{a_n \mid n \in \mathbb{N}\}$. S is nonempty because $a_1 \in S$. S is bounded above by definition of boundedness of a sequence. Thus S has a least upper bound, which we will call L.

We claim that $\lim_{n \to \infty} a_n = L$. Fix $\epsilon > 0$. $L - \epsilon$ is not an upper bound for S, so there is an element of S greater than $L - \epsilon$, so there is an N such that $L - \epsilon < a_N$. Now for any $n > N$ we have $L - \epsilon < a_N \le a_n \le L < L + \epsilon$. The first inequality is because a_N is chosen to witness $L - \epsilon$ not being an upper bound, the second holds because the sequence is nondecreasing, the third holds because L is an upper bound for S and the fourth is obvious. We see that $|a_n - L| < \epsilon$ for $n > N$ so we are done.

Corollary: A nonincreasing sequence which is bounded below converges. If $\{a_n\}$ is nonincreasing and bounded below, then $\{-a_n\}$ is nondecreasing and bounded above...

Corollary: A monotone bounded sequence conveerges. The only additional detail you need to notice is that a nondecreasing sequence bounded

above or a nonincreasing sequence bounded below is actually bounded both above and below (the other bound in each case can be taken to be its first term).

7 Problems 12 and 14, chapter 6

Problem 12a, chapter 6: STAR Prove that if f is continuous at L and $\lim_{x \to a} g(x) = L$, then $\lim_{x \to a} f(g(x)) = f(L)$. Do not use theorems: prove this straight from the definition of limit and algebra/inequalities). (This rules out the answers using the theorem from the chapter which I accepted on the homework).

Fix $\epsilon_0 > 0$.

Because f is continuous at L, we know that $\lim_{x \to a} f(x) = f(L)$.

We know that there is a δ_1 such that for any u, $|f(u) - f(L)| < \epsilon$ if $|u - L| < \delta_1$.

Because $\lim_{x \to a} g(x) = L$, we know that there is a δ_2 such that if $0 < |x - a| < \delta_2$, then $|g(x) - L| < \delta_1$ (yes, we are using δ_1 as an epsilon here).

Now let x be chosen arbitrarily and assum $0 < |x - a| < \delta_2$. It follows that $|g(x) - L| < \delta_1$, from which it follows by choice of δ_1 and setting $u = g(x)$ that $|f(g(x)) - f(L)| < \epsilon_0$, which is what we need.

Problem 14, chapter 6: Suppose that g and h are continuous at a and $g(a) = h(a)$. Define $f(x)$ as $g(x)$ for $x \leq a$ and $h(x)$ for $x \geq a$. Show that f is continuous at a.

Fix $\epsilon_0 > 0$.

Because g is continuous at a, there is δ_1 such that for any x, if $|x - a| < \delta_1$ then $|g(x) - g(a)| < \epsilon_0$.

Because h is continuous at a, there is δ_2 such that for any x, if $|x - a| < \delta_2$ then $|h(x) - h(a)| < \epsilon_0$.

Let $\delta_0 = \min(\delta_1, \delta_2)$. Let x be chosen arbitrarily; assume $0 < |x - a| < \delta_0$.

There are two cases: either $x < a$ or $x > a$.

If $x < a$, note that $0 < |x - a| < \delta_0 \leq \delta_1$, so $|g(x) - g(a)| < \epsilon_0$. We have $g(x) = f(x)$ because $x < a$, and $g(a) = f(a)$, so $|f(x) - f(a)| < \epsilon_0$.

If $x > a$, note that $0 < |x - a| < \delta_0 \leq \delta_2$, so $|h(x) - h(a)| < \epsilon_0$. We have $h(x) = f(x)$ because $x > a$, and $h(a) = f(a)$, so $|f(x) - f(a)| < \epsilon_0$

Thus in either case, if $0 < |x - a| < \delta_0$ then $|f(x) - f(a)| < \epsilon_0$, so $\lim_{x \to a} f(x) = f(a)$, that is, f is continuous at a.

8   More about Sequences

Theorem 1, chapter 22, Part I: STAR Suppose $\lim_{x \to c} f(x) = L$ and suppose a is a sequence such that each $f(a_i)$ is defined, each $a_i \neq c$, and $\lim_{i \to \infty} a_i = c$. Then $\lim_{i \to \infty} f(a_i) = L$.

Proof: This is closely related to chapter 6 problem 12.

Fix $\epsilon_0 > 0$.

We know that there is a δ_1 such that for any u, if $0 < |u - c| < \delta_1$, then $|f(u) - L| < \epsilon$, by the assumption about the limit of f at c.

We know further that there is an N_1 such that for any $i > N_1$ we have $|a_i - c| < \delta_1$, by the assumption about the limit of the sequence $\{a_i\}$, with ϵ set equal to δ_1 (the same sort of thing that happened in the proof of problem 12 chapter 6).

Now choose i arbitrarily. Assume $i > N_1$. It follows that $|a_i - c| < \delta_1$. From this it follows that $|f(a_i) - L| < \epsilon_0$.

So we have shown that for an arbitrarily chosen ϵ_0, we can find an N_1 such that if $i > N_1$ then $|f(a_i) - L| < \epsilon_0$, that is, $\lim_{i \to \infty} f(a_i) = L$.

Theorem 1, chapter 22, Part II: We show that if f is defined on an interval including c (except possibly at c) and for every sequence a such that $f(a_i)$ is always defined, no a_i is equal to c and $\lim_{i \to \infty} a_i = c$ we also have $\lim_{x \to c} f(a_i) = L$, then it follows that $\lim_{x \to c} f(x) = L$ (a fact about convergence of a function can be deduced from facts about convergence of series).

We argue indirectly: we show that if f is defined on an interval around c (except possibly at c) and $\lim_{x \to c} f(x) \neq L$ we can construct a sequence $\{a_i\}$ such that $\lim_{i \to \infty} a_i = c$ but $\lim_{i \to \infty} f(a_i) \neq L$.

$\lim_{x \to c} f(x) \neq L$ means $\neg$(for each $\epsilon > 0$ there is a $\delta > 0$ such that for all x if $0 < |x - c| < \delta$ then $|f(x) - L| < \epsilon$) which transforms using de Morgan's laws for quantifiers to "for some $\epsilon > 0$ for all $\delta > 0$ there is x such that $0 < |x - c| < \delta$ and $|f(x) - L| \geq \epsilon$" [the assumption that f is defined close enough to c rules out the alternative of this failing to be true simply by f not having a value at points arbitrairly close to c].

Choose a fixed ϵ_0 for which this is true. For each natural number i, there is a number a_i such that $0 < |a_i - c| < \frac{1}{i}$ and $|f(a_i) - L| \geq \epsilon_0$. This is obtained from the stated property of ϵ_0 by setting $\delta = \frac{1}{i}$. We claim that the sequence of numbers a_i we have just selected converges to c and has $f(a_i)$ not converging to L.

For any $\epsilon > 0$, there is an $N > \frac{1}{\epsilon}$, and for each $i > N$ we have $|a_i - c| < \frac{1}{i} < \frac{1}{N} < \epsilon$. So $\lim_{i \to \infty} a_i = c$.

For the specific value ϵ_0, $|f(a_i) - L| \geq \epsilon_0$ for all i, so there cannot be an N such that for all $i > N$, $|f(a_i) - L| < \epsilon_0$, so it is not the case that $\lim_{i \to \infty} f(a_i) = L$.

9 Notes from the last day

Another approach to problem 12 chapter 8: Suppose that A and B are nonempty sets and for any $a \in A$ and $b \in B$ we have $a \leq b$.

We aim to prove that $\sup(A) \leq \inf(B)$.

Suppose for the sake of a contradiction that $\inf(B) < \sup(A)$. Since $\inf(B)$ is less than the least upper bound of A, it is not an upper bound of A, so we can select an a_1 such that $\inf(B) < a_1 \leq \sup(A)$. Now because a_1 is greater than the greatest lower bound of B, a_1 is not a lower bound of B, so we can select $b_1 \in B$ such that $\inf(B) \leq b_1 < a_1 \leq \sup(A)$. But this is a contradiction, as our assumptions require $a_1 \leq b_1$.

chapter 8, problem 5: 1. Suppose $y - x > 1$. Show that there is an integer n such that $x < n < y$.

Any set of integers which is bounded above has a largest element. In particular, there is a largest integer $\leq x$. Call this integer L. We have $x < L + 1$, because $L + 1 \leq x$ is ruled out by L being the largest integer $\leq x$. We have $L + 1 \leq x + 1$ by adding 1 to both

sides of $L \leq x$. We have $x + 1 < y$ by adding 1 to both sides of the assumption $y - x > 1$. So we have $x < L + 1 < y$, and $L + 1$ is the desired integer n.

2. Suppose $x < y$. Then there is a rational number r with $x < r < y$ Since $y - x > 0$, there is a natural number N such that $N > \frac{1}{y-x}$ (because the natural numbers are unbounded: remember I said we would need that result in this problem!) and this implies $y - x > \frac{1}{N}$. From this it follows that $Ny - Nx > 1$, so by the first part there is an integer I such that $Nx < I < Ny$, so $x < \frac{I}{N} < y$, and $\frac{I}{N}$ is a suitable rational number r.

3. Suppose that r and s are rationals and $r < s$. Then there is an irrational x such that $r < x < s$.

   By problem 2 chapter 4, for any t with $0 < t < 1$ we have $a < (1-t)a + tb < b$, so we have $r < (1 - \frac{1}{\sqrt{2}})r + \frac{1}{\sqrt{2}}s < s$, and if r and s are rational, $x = (1 - \frac{1}{\sqrt{2}})r + \frac{1}{\sqrt{2}}s$ will be irrational (because we can show that from any rational u between r and s we can compute the corresponding t as $\frac{u-r}{s-r}$, and this will be rational if u is rational.)

4. Suppose that $x < y$ are numbers. Then there is an irrational z such that $x < z < y$.

   By part 2, we know that there is a rational r such that $x < r < y$. By part 2 we know that there is a rational s such that $r < s < y$. By part 3, we know that there is an irrational z with $r < z < s$. So we have $x < r < z < s < y$, so we have $x < z < y$ for an irrational z.

definition of convergence: A sequence $\{a_n\}$ converges iff there is an L such that $\lim_{i \to \infty} a_i = L$: a sequence is convergent iff it has a limit.

a convergent sequence must be bounded: STAR

Fix $\epsilon > 0$ (it really doesnt matter which ϵ we pick).

there is an N such that for all $i > N$ we have $|a_i - L| < \epsilon$, that is $L - \epsilon < a_i < L + \epsilon$.

So $L + \epsilon$ is an upper bound for all terms past the Nth term and of course there are only finitely many terms before the Nth term:

$\max(\max_{i \leq N} a_i, L + \epsilon)$ is an upper bound for all the a_i's, and similarly $\min(\min_{i \leq N} a_i, L - \epsilon)$ is a lower bound. The crucial point here is that of course a finite set of numbers has an upper bound and in fact a maximum element. The notation $\max_{i \leq N} a_i$ stands for the largest of the numbers $a_1, \ldots, a_N$. The notation $\min_{i \leq N} a_i$ stands for the smallest of the numbers $a_1, \ldots, a_N$.

The problem which says that if $\lim_{i \to \infty} a_i = 0$ and some $a_n > 0$ then there must be a maximum term in the sequence has a very similar proof: think about the fact that for some N, for all $i > N$, $|a_i| < \frac{a_n}{2}$ (so in particular all a_i's for $i > N$ are less than a_n) is what is needed, along with the fact that any finite set of numbers has a maximum element.

Don't confuse implications with their converses: It is NOT true that a bounded sequence must converge. For example $a_n = (-1)^{n+1}$ does not converge $(1, -1, 1, -1, \ldots)$.

It is also NOT true that a convergent sequence is nondecreasing and bounded above : we have the converse as a theorem (any nondecreasing sequence which is bounded above converges) but don't confuse an implication with its converse. The sequence $a_n = (-1)^{n+1}\frac{1}{n}$ is neither increasing nor decreasing but converges to 0 $(1, -\frac{1}{2}, \frac{1}{3}, -\frac{1}{4} \ldots)$.

Definition of a subsequence: A subsequence of a sequence is obtained by selecting elements from the parent sequence in increasing order of index. The formal definition is as follows:

$\{b_i\}$ is a subsequence of $\{a_i\}$ iff there is a strictly increasing sequence of natural numbers s_i such that for each i, $b_i = a_{s_i}$.

For example, if $a_i = \frac{1}{i}$ we have the sequence $1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5} \ldots$. If $s_i = 2i-1$ $(1, 3, 5 \ldots)$ the sequence b_i is $a_1, a_3, a_5 \ldots$, in this case $1, \frac{1}{3}, \frac{1}{5} \ldots$.Our aim

A subsequence of a convergent subsequence is convergent: The statement is stronger than this. If a sequence has limit L, then all of its subsequences have the same limit L.

Assume that $\lim_{i \to \infty} a_i = L$. Our aim is to show that for any strictly increasing sequence s_i of natural numbers, $\lim_{i \to \infty} a_{s_i} = L$. Since every subsequence of $\{a_i\}$ can be written in the form a_{s_i}, this is enough.

Fix an $\epsilon > 0$. We know that there is N such that for all $i > N$, we have $|a_i - L| < N$. But also we have $s_i \geq i > N$, so we also have

$|a_{s_i} - L| < \epsilon$ for all i greater than the same N.

It should be evident that for any strictly increasing sequence s_i of natural numbers, $s_i \geq i$ for all i: the sequence must go up by at least one at each step. A formal proof would use math induction.

Bolzano-Weierstrass theorem: We claim that any bounded sequence has a convergent subsequence.

We say that a sequence $\{a_i\}$ has a high point at k iff for all $n > k$, $a_n < a_k$: every term after a_k is strictly less than a_k.

Assume that $\{a_i\}$ is a bounded sequence.

Either there are infinitely many high points of $\{a_i\}$ or there are not.

If there are infinitely many high points of $\{a_i\}$, we define a subsequence $\{a_{s_i}\}$ as follows: s_1 is the smallest natural number k such that a has a high point at k, and s_{i+1} (for each i) is the smallest $k > s_i$ such that $\{a_i\}$ has a high point at k. Now observe that for any $i < j$, $a_{s_i} > a_{s_j}$, because there is a high point of $\{a_i\}$ at s_i. Thus the subsequence $\{a_{s_i}\}$ is strictly decreasing. It is bounded below because a is bounded below, so it converges by the monotone convergence theorem.

If there are only finitely many high points of $\{a_i\}$, we define a subsequence a_{s_i} as follows: let s_1 be $b + 1$, where b is the largest number such that $\{a_i\}$ has a high point at b. Let s_{i+1} be the least $k > s_i$ such that $a_k \geq a_{s_i}$. We can be sure that s_{i+1} is defined, because we know that $\{a_i\}$ does not have a high point at s_i (since it has no high point at or above s_1). The subsequence $\{a_{s_i}\}$ is obviously nondecreasing from the way it is defined, and it is bounded above because $\{a_i\}$ is bounded above, so it converges by the monotone convergence theorem.

We reiterate that it is NOT true that every bounded sequence is convergent. $1, -1, 1, -1, 1, -1, \ldots$ is an obvious counterexample – and its convergent subsequences are also obvious (constant sequences converging to 1 and -1).

Prove the extreme value theorem using the Bolzano-Weierstrass theorem: We have no time to see why this theorem is important; we slip in one quick example.

We prove (as we already proved above using a different method) that if a function f is continuous on a closed interval $[a, b]$, then the range

of f is bounded above.

Suppose for the sake of a contradiction that f is continuous on $[a,b]$ and the range of f is not bounded above.

Then we can select a sequence $\{x_i\}$ with each $x_i \in [a,b]$ and $f(x_i) > i$ for each i (we can do this because f is assumed to take on unboundedly large values).

It follows that there is a strictly increasing sequence of natural numbers s_i such that $\{x_{s_i}\}$ converges to a limit c. $c \in [a,b]$ must be true (this was a problem in your homework; I prove it below). Now it follows by theorem 1 part 1 chapter 22 that $\lim_{i \to \infty} f(x_{s_i}) = f(c)$. This means that for any $\epsilon > 0$ there is an N such that for all $i > N$, $|f(x_{s_i}) - f(c)| < \epsilon$, so for all $i > N$, $f(x_{s_i}) < f(c) + \epsilon$. But observe that for any $i > f(c) + \epsilon$ we have $f(x_{s_i}) > s_i \geq i > f(c) + \epsilon$, which is a contradiction (if we consider i greater than the maximum of N and $f(c) + \epsilon$ we find that $f(x_{s_i})$ must be both greater than and less than $f(c) + \epsilon$). If you are observant, you may notice that the contradiction can be obtained much faster: a convergent sequence is bounded, and $\{f(x_{s_i})\}$ both has to converge to $f(c)$ and is obviously unbounded.

A convergent sequence of numbers in a closed interval has limit in that interval:

Suppose that $\{c_i\}$ is a convergent sequence such that $c_i \in [a,b]$ for each i. We claim that $\lim_{i \to \infty} c_i$, which we will call c, must belong to $[a,b]$.

If it doesn't then it must either be greater than b or less than a. If $c > b$, then we can find an N such that for all $i > N$, $|c_i - c| < c - b$, so in particular $c_i > c - (c - b) = b$, which contradicts our assumptions. Similarly, if $c < a$, then we can find an N such that for all $i > N$, $|c_i - c| < a - c$, so in particular $c_i < c - (a - c) = a$, which contradicts our assumptions.