MATHEMATICS 300 — SPRING 2015

Introduction to Mathematical Reasoning

H. J. Sussmann

INSTRUCTOR'S NOTES

LECTURES NOS. 2, 3, AND 4

Contents

1 The real number system

We would now like to start our systematic study of number systems, by looking at the real numbers. It turns out that the main thing that makes the real numbers interesting is the operations (addition, subtraction, multiplication and division) on them, and the predicate "less than."

So, before we start our study of $\mathbb{R}$, let us say a few words about operations and predicates.

1.1 Operations and relations

An **operation** on a set S is a rule that can be applied to one or several members of S to produce another member of S. A **one-argument operation** on S produces, for every member x of S, a member (the result of applying the operation to x) that would be called $f(x)$, or some expression involving x. Examples of one-argument operations are:

1. **minus** (or **negative of**), i.e., the operation on $\mathbb{R}$ that for every $x \in \mathbb{R}$ produces the number $-x$, called "minus x", or "the negative of x";

2. **inverse of**, i.e., the operation on $\mathbb{R}$ that for every $x \in \mathbb{R}$ such that $x \neq 0$ produces the number x^{-1} (also written $\frac{1}{x}$), called the **multiplicative inverse** of x (this operation is *partially defined*, because x^{-1} is not defined for all $x \in \mathbb{R}$, but only for those $x \in \mathbb{R}$ other than $x = 0$);

3. **absolute value**, i.e., the operation on $\mathbb{R}$ that for every $x \in \mathbb{R}$ produces the number $|x|$, called the **absolute value** of x;

4. **square**, i.e., the operation on $\mathbb{R}$ that for every $x \in \mathbb{R}$ produces the number x^2, called the **square** of x,

5. **square root**, i.e., the partially defined operation on $\mathbb{R}$ that for every $x \in \mathbb{R}$ such that $x \geq 0$ produces the number $\sqrt{x}$ (defined to be the unique $y \in \mathbb{R}$ such that $y^2 = x$ and $y \geq 0$), called the **square root** of x.

A **two-argument operation** (more commonly called **binary operation**) on S produces, for every pair (x, y) of members of S, a member (the result of applying the operation to x and y) that would be called $f(x, y)$, or some expression involving x and y. Examples of binary operations are:

1. **addition** of real numbers, i.e., the operation that for every $x \in \mathbb{R}$ and every $y \in \mathbb{R}$ produces the number $x+y$, called "x plus y", or "the sum of x and y";

2. **multiplication** of real numbers, i.e., the operation that for every $x \in \mathbb{R}$ and every $y \in \mathbb{R}$ produces the number $x \times y$ (also written $x.y$, or xy), called "x times y", or "the product of x and y");

3. **subtraction** of real numbers, i.e., the operation that for every $x \in \mathbb{R}$ and every $y \in \mathbb{R}$ produces the number $x - y$, called "x minus y", or "the difference of x and y";

4. **division** of real numbers, i.e., the operation that for every $x \in \mathbb{R}$ and every $y \in \mathbb{R}$ such that $y \neq 0$ produces the number $x \div y$ (also written x/y, or $\frac{x}{y}$), called "x over y", or "x divided by y", or "the quotient of x over y".

In principle, there can also be **three-argument operations** (more commonly called **ternary operations**), **four-argument operations**, and in fact n-argument operations for any natural number n. For example:

1. The **average of three real numbers** is the operation that, for any three real numbers x, y, z, produces the number $\frac{x+y+z}{3}$, called the **average** of x, y, and z. So "average of three real numbers" is a three-argument operation on $\mathbb{R}$.

A **relation** (or **predicate**) on a set S is a rule that can be applied to one or several members of S to produce a truth value, "true" or "false". A **one-argument predicate** on S produces, for every member x of S, a true-or-false value. For example, "positive" is a one-argument predicate on $\mathbb{R}$: for each $x \in \mathbb{R}$, "x is positive" is true—in which case we write "$x > 0$"—or false, in which case we write "$\sim x > 0$". For a second example, "is an integer" can be regarded as a one-argument predicate on $\mathbb{R}$: for each $x \in \mathbb{R}$, "x is an integer" is true—in which case we write "$x \in \mathbb{Z}$"—or false, in which case we write "$\sim x \in \mathbb{Z}$" or "$x \notin \mathbb{Z}$".

A **two-argument predicate**, or **two-argument relation**, or **binary relation**, on S produces, for every pair (x, y) of members of S, a true-or-false value. For example, "less than" is a binary relation on $\mathbb{R}$: for each $x \in \mathbb{R}$ and each $y \in \mathbb{R}$, "x is less than y" is either true—in which case we write "$x < y$"—or false, in which case we write "$\sim x < y$". For a second example, "is equal to" is a binary relation on $\mathbb{R}$ or, more generally, on any set S: for any two members x, y of S, "x is equal to y" (or "x is the same as y", or "x is y") is either true—in which case we write "$x = y$"—or false, in which case we write "$\sim x = y$" or "$x \neq y$".

There are also three-argument predicates, four-argument predicates, and predicates with any number of arguments. For example, "x is smaller than the sum of y and z" is a three-argument predicate on $\mathbb{R}$: for any three real numbers x, y, z, either it is true that x is smaller than the sum of y and z—in which case we write "$x < y + z$"—or it is false, in which case we write "$\sim x < y + z$".

Remark 1. You are probably familiar with the concept of "function", that will be discussed later in the course. Using this concept, we can say that an n-argument operation on a set S is a function of n variables on S that takes values in S, and an n-argument predicate on a set S is a function of n variables on S that takes values in the set $\{true, false\}$. □

1.2 The basic concepts of real number theory

To study of the real numbers, we will want to analyze several concepts and properties associated with them, such as, for example:

1. the operations of addition, subtraction, multiplication and division of real numbers,
2. the number zero,
3. the number one,
4. the order relation ("$<$"),
5. the square of a real number and, more generally, powers of all kinds,
6. the absolute value of a real number,
7. integers and natural numbers.

The approach we will use is to start with some **basic concepts and properties**, and then **define** all the other ones. The basic concepts and properties will be represented by symbols, such as 0, 1, $+$, $\times$, $<$, and all other concepts of the theory will be defined in terms of them, and new symbols will be introduced to represent them. So, for example, "2" is not one of the basic concepts, so we will have to define "2". (This is easy: we will just define 2 to be $1+1$.) And "absolute value" is not one of the basic concepts either, so we will have to define what "absolute value" means.

Here is the list of the basic concepts of real number theory:

1. the numbers 0 (zero), and 1 (one),
2. the binary operations of addition, subtraction, multiplication and division:
    i. *addition* produces, for any two real numbers x, y, a real number $x+y$, called the *sum* of x and y,
    ii. *subtraction* produces, for any two real numbers x, y, a real number $x-y$, called the *difference* of x and y, (that is "x minus y"),
    iii. *multiplication* produces, for any two real numbers x, y, a real number xy, called the *product* of x and y; we also write[1] $x.y$, or $x \times y$, for the product xy.
    iv. *division* produces, for any two real numbers x, y such that $y \neq 0$, a number $x \div y$ (also written x/y, or $\frac{x}{y}$), called the *quotient* of x over y. Division is a *partially defined* operation, because the quotient $x \div y$ does not make sense for all possible real numbers x and y: $x \div y$ only makes sense when y is not equal to zero.
3. the *order relation* $<$ ("less than"): for any two real numbers x and y, x is either less than y or not. We write "$x < y$" to indicate that x is less than y.

Starting with these basic concepts, we will want to *define* all other concepts and properties of interest in the theory.

For example: what is the "absolute value" of a real number? Since "absolute value" is not one of our basic concepts, we have to *define* absolute value in terms of the basic concepts. And we cannot define the absolute

[1] This is especially useful when we are dealing with specific numbers represented by "numerals", i.e., symbols such as 23 or 3.72. If we want to write the product of 23 and 45, it is better not to write 2345, because this is the name of the number two-thousand three hundred and forty-five, rather than the product of 23 and 45. So it is much better to write 23×45.

value of a real number to be the "magnitude" of the number, because "magnitude" is not one of the basic concepts, so saying that "the absolute value of a real number is its magnitude" is meaningless, because we do not know what "magnitude" means[2].

Here is a correct way to define "absolute value".

Definition 1. Given a real number x, the absolute value of x is the number $|x|$ defined as follows:

$$\begin{aligned} |x| &= x \quad \text{if } 0 < x\,, \\ |x| &= -x \quad \text{if } x < 0\,, \\ |x| &= 0 \quad \text{if } x = 0\,. \end{aligned}$$

□

Example 1.

1. $|5| = 5$, because $0 < 5$, so (1.1) applies.
2. $|-5| = 5$, because $-5 < 0$, so (1.1) applies, and $|-5| = -(-5) = 5$.
3. $|0| = 0$. □

1.3 The basic facts of real number theory. Part I: the field axioms

Now that we have described the basic symbols of real number theory, we list the basic properties of the concepts represented by those symbols. These basic properties are the *axioms* for the real numbers, that is, the facts that we take from granted and use as the starting point of the development of the theory. **Everything else has to be proved.**

We divide the list of axioms into two parts. First, in this subsection, we present the axioms about 0, 1, $+$, $\times$, $-$, and $\div$. And then, in a later section, we will list the axioms involving $<$.

The axioms about 0, 1, $+$, $\times$, $-$, and $\div$ are called the field axioms, because any system of "numbers" in which there are special "numbers" 0

[2]Based on my own experience of teaching Math 300 many times, I can predict that, when asked to define "absolute value" in one of the midterms or the final exam, many students are going to write "the absolute value of a real number is its magnitude". **Please do not do that!**

and 1, and operations $+$, $\times$, $-$, and $\div$ that obey these axioms is called a field. (We will see later in the course examples of fields other than $\mathbb{R}$.)

And, before we actually list the axioms, we have to say a few words about equality ("=") and inequality ("$\neq$"), because these concepts will appear in the axioms. So we digress a little bit and talk about equality.

1.4 A detour: equality and inequality

If x and y are any objects (numbers, sets, people, whatever), we write "$x = y$" (and read this as "x is equal to y", or "x and y are equal") to indicate that x and y are the same object. And we write "$x \neq y$" to indicate that x and y are not the same, so "$x \neq y$" means exactly the same thing as "$\sim x = y$".

Equality obeys the following laws (called "equality axioms"):

The axioms for equality

EA1. (reflexive law of equality) If x is any object, then $x = x$.

EA2. (symmetry law of equality) If x, y are any objects, and $x = y$, then $y = x$.

EA3. (transitivity law of equality) If x, y, z are any objects, $x = y$, and $y = z$, then $x = z$.

In addition, equality satisfies the following "substitution of equals for equals" rule, that can be used in proofs[3]:

[3] The rule talks about "terms" and "sentences", so let me say a few words about that. Terms and sentences will be discussed in detail later. At this point, all you need to know is that a term is an expression that is the name of an object, and a sentence is an expression that makes an assertion that can be true or false. For example, "1", "1 + 1", "2 + 3", "(7.43 + 22.04) × 96", "Mt. Everest", "Lady Gaga", and "The man who came to dinner yesterday evening but didn't stay very long because he had another engagement" are terms, and "2 + 2 = 4", "Mt. Everest is taller than Mt. McKinley", "Lady Gaga sang together with Tony Bennett", and "The man who came to dinner yesterday evening but didn't stay very long because he had another engagement told us that he had enjoyed the dinner very much" are sentences.

The substitution of equals for equals rule (Rule SEE)

If

a. S is a statement containing, once or several times, a term T

b. U is another term,

c. we have $U = T$ or $T = U$ in an earlier step of our proof,

d. we have S in an earlier step of our proof,

then we can assert, in a new step of our proof, a sentence obtained from S by substituting for T the term U, in some or all the occurrences of T in S.

Example 2. If we have $2 = 1 + 1$ and $2 + 1 = 3$ in earlier steps, then we can write $(1 + 1) + 1 = 3$. (Here, S is the sentence "$2 + 1 = 3$", T is the term "2", and U is the term "$1 + 1$". Substitution of U for T is S yields "$(1 + 1) + 1 = 3$.) □

Example 3. If we have $2 = 1+1$ and $2+2 = 4$ in earlier steps, we can write $(1+1)+2 = 4$, or $2+(1+1) = 4$, or $(1+1)+(1+1) = 4$. (That is, we can substitute "$1+1$" for "2" in the first of the two 2s that occur in $2+2 = 4$, or in the second one, or in both.) □

And now we are ready to go back to the discussion of the real number axioms.

1.5 The list of the field axioms

FA1. (closure laws) If x, y are real numbers, then

- FA1.a $x + y$ is a real number,
- FA1.b $x - y$ is a real number,
- FA1.c xy is a real number,
- FA1.d if $y \neq 0$ then $x \div y$ is a real number.

FA2. (associative law of addition) If x, y, z are real numbers, then

$$(x+y)+z = x+(y+z)\,.$$

FA3. (commutative law of addition) If x, y are real numbers, then

$$x+y = y+x\,.$$

FA4. (associative law of multiplication) If x, y, z are real numbers, then $(xy)z = x(yz)$.

FA5. (commutative law of multiplication) If x, y are real numbers, then $xy = yx$.

FA6. (distributive law of multiplication with respect to addition) If x, y, z are real numbers, then $x(y+z) = xy + xz$.

FA7. (subtraction axiom) If x, y are real numbers, then $(x-y)+y = x$.

FA8. (division axiom) If x, y are real numbers, and $y \neq 0$, then $(x \div y).y = x$.

FA9. (additive identity law) If x is a real number, then $x+0 = x$.

FA10. (multiplicative identity law) If x is a real number, then $x \times 1 = x$.

FA11[4]. $0 \neq 1$.

2 A brief detour into Logic: formal language, quantifiers, "and", "or", and "implies"

If you look at any of the 11 field axioms listed in the previous subsection, you will see that their statement is made in a mixture of ordinary language and formulas. For example, Axiom FA6 says: "If x, y, z are real numbers, then $x(y+z) = xy+xz$". This uses the formula $x(y+z) = xy+xz$, and also English words.

It turns out that mathematicians, and logicians, have invented "formal languages", in which you can say *everything* with formulas, without using any words. It is going to be very important for us to learn formal language, and to be able to translate from English to formal language and back.

Remark 2. Why is this important? There are several reasons, and we will talk about them later. For the moment, let me give you just two reasons:

- ***In formal language, you are obliged to be absolutely precise.*** If you are saying something in English, and you cannot translate it into formal language, it means that you really do not know what it is exactly that you are trying to say. For example, the statements "25 is a small number", or "x is a number", cannot be translated into formal language, and this is an indication that you have to think some more, figure out exactly what you are trying to say, and once you know precisely what it is that you want to say, then you will be able to say it in formal language.

  Why is "42 is a small number" not precise? The answer is, simply, that in Mathematics there is no such things as a "small number". Smallness depends very much on the context. If you are talking about the number of people who watched the President's State of the Union speech on TV, then 42 is a very small number, since one would expect that millions of people would have watched the speech. But if you are talking about the number of candidates running for the Republican presidential nomination, then 42 is a huge number.

  Similarly, "x is a number" is not a precise statement. (See Lecture 1, Page 9.) You can say "x is a real number" in formal language, by saying "$x \in \mathbb{R}$". Or you can say "x is an integer", by saying "$x \in \mathbb{Z}$". So one way for you to realize that you are not supposed to say "x is a number", is to try to say it in formal language and see that you cannot do it.

- ***Formal language is a truly internationa language.*** The formulas of formal mathematical language are the same in any language. So when a Chinese mathematician publishbes a paper in Chinese in a Chinese journal, every mathematician in the world can read the formulas.

□

2.1 The quantifiers: "∀" and "∃"

We now begin our discussion of the symbols of formal language by talking about the two **quantifiers**: existential and universal. (We already talked about existential quantifiers in Lecture 1, pages 10-11.)

The symbols ∃ and ∀ are the ***quantifier symbols***:

- " ∃ " is the ***existential quantifier symbol***,

and

- " ∀ " is the ***universal quantifier symbol***.

Using these symbols, we can form ***quantifiers***.

- An ***existential quantifier*** is an expression "$(\exists x)$" or "$(\exists x \in S)$" (if S is a set).
  - "$(\exists x)$" is an ***unrestricted existential quantifier***,

  and

  - ''$(\exists x \in S)$" is a ***restricted existential quantifier***.

Similarly,

- a ***universal quantifier*** is an expression "$(\forall x)$" or "$(\forall x \in S)$" (if S is a set).
  - "$(\forall x)$" is an ***unrestricted universal quantifier***,

  and

  - ''$(\forall x \in S)$" is a ***restricted universal quantifier***.

Quantifiers are read as follows:

1. "$(\exists x)$" is read as
   - "there exists x such that"

   or

   - "for some x"

   or

   - "it is possible to pick x such that".
2. "$(\exists x \in S)$" is read as

- "there exists x belonging to S such that"

or

- "there exists a member x of S such that"

or

- "for some x in S"

or

- "it is possible to pick x in S such that"

or

- "it is possible to pick a member x of S such that"

3. "$(\forall x)$" is read as

- "for all x"

or

- "for every x "

or

- "given any x"

or

- "no matter who x is"

4. "$(\forall x \in S)$" is read as

- "for all x in S"

or

- "for every x in S "

or

- "given any x in S"

or

- “no matter who x in S is”

or

- “for all members x of S”

or

- “for every member x of S ”

or

- “given any member x of S”

or

- “for all x belonging to S”

or

- “for every x belonging to S ”

or

- “given any x belonging to S”.

Example 4. The sentence

$$(\exists x \in \mathbb{R})x^2 = 2 \tag{2.1}$$

could be read as “there exists an x belonging to the set of real numbers such that $x^2 = 2$”, but a much better way to read it is: “there exists a real number x such that $x^2 = 2$”. You can also read it as: “it is possible to pick a real number x such that $x^2 = 2$”. □

Example 5. The sentence[5]

$$(\forall x \in \mathbb{R})x^2 > 0 \tag{2.2}$$

[5] I am not saying that this sentence is true. I am just talking about the sentence. Actually, Sentence 2.2 is **false**.

could be read as "for every x belonging to the set of real numbers, x^2 is positive", but a much better reading is "for every real number x, x^2 is positive".

And an even better reading is: "the square of every real number is positive".

And another very nice way to read this sentence is "if x is an arbitrary[6] real number, then x-square is positive."

Finally, another nice way to read it is "the square of an arbitrary real number is positive". □

Example 6. The best ways to read the sentence

$$(\forall x \in \mathbb{R})x^2 \geq 0 \tag{2.3}$$

are

- "the square of every real number is nonnegative",
- "if x is an arbitrary real number, then x-square is nonnegative",

and

- "the square of an arbitrary real number is nonnegative". □

Remark 3. Sentence (2.1) is true, because there is a real number x whose square is 2. (Reason: take $x = \sqrt{2}$.)

Sentence (2.2) is false, because if you take $x = 0$ then "$x^2 > 0$" is not true, so it is not true that for *every* real number x the square of x is > 0.

Sentence (2.3) is true. □

2.2 The connectives "∧" (meaning "and"), "∨" (meaning "or") , and "⟹" (meaning "implies")

Conjunction: The symbol "∧" is the ***conjunction symbol***, and means "and". So, for example, if P is the sentence "today is Friday" and Q is the sentence "tomorrow is Saturday", then "$P \wedge Q$" stands for the sentence "today is Friday and tomorrow is Saturday". A sentence of the form $P \wedge Q$ is a **conjunction**. And, in a conjunction $P \wedge Q$, the sentences P, Q are the **conjuncts**.

[6] The meaning of "arbitrary" will be explained in page 18.

Disjunction: The symbol "$\vee$" is the ***disjunction symbol***, and means "or". So, for example, if P is the sentence "today is Friday" and Q is the sentence "today is Saturday", then "$P \vee Q$" stands for the sentence "today is Friday or today is Saturday". A sentence of the form $P \vee Q$ is a **disjunction**. And, in a disjunction $P \vee Q$, the sentences P, Q are the **disjuncts**.

Implication: The symbol "$\Longrightarrow$" is the ***implication symbol***, and means "implies". A sentence "$P \Longrightarrow Q$" is read as "P implies Q", or "If P then Q". So, for example, if P is the sentence "today is Friday" and Q is the sentence "tomorrow is Saturday", then "$P \Longrightarrow Q$" stands for the sentence "If today is Friday then tomorrow is Saturday". A sentence of the form $P \Longrightarrow Q$ is an **implication**, or a **conditional sentence**. And, in a conditional sentence $P \Longrightarrow Q$, P is the **premiss** (or **antecedent**), and Q is the **conclusion** (or **consequent**.

Remark 4. Notice that "$\wedge$" and "$\Longrightarrow$" are very different. For example, the sentence "today is Friday and tomorrow is Saturday" is true only if today is Friday. On the other hand, the sentence "If today is Friday then tomorrow is Saturday" is true no matter what day it is today. (Think of "If today is Friday then tomorrow is Saturday" as meaning "If today was Friday then tomorrow would be Saturday". This is always true, even is today happens to be Tuesday. If you are not convinced, wait. Implication will ve discussed later.) $\square$

3 The field axioms restated in formal language

Using universal quantifiers, conjunctions and implications, let us restate all the field axioms into formal language. Here they are:

The field axioms for $\mathbb{R}$.

FA1. a $(\forall x \in \mathbb{R})(\forall y \in \mathbb{R})x + y \in \mathbb{R}$,

b $(\forall x \in \mathbb{R})(\forall y \in \mathbb{R})x - y \in \mathbb{R}$,

c $(\forall x \in \mathbb{R})(\forall y \in \mathbb{R})xy \in \mathbb{R}$,

d $(\forall x \in \mathbb{R})(\forall y \in \mathbb{R})(y \neq 0 \Longrightarrow x \div y \in \mathbb{R})$.

FA2. $(\forall x \in \mathbb{R})(\forall y \in \mathbb{R})(\forall z \in \mathbb{R})(x + y) + z = x + (y + z)$.

FA3. $(\forall x \in \mathbb{R})(\forall y \in \mathbb{R})x + y = y + x$.

FA4. $(\forall x \in \mathbb{R})(\forall y \in \mathbb{R})(\forall z \in \mathbb{R})(xy)z = x(yz)$.

FA5. $(\forall x \in \mathbb{R})(\forall y \in \mathbb{R})xy = yx$.

FA6. $(\forall x \in \mathbb{R})(\forall y \in \mathbb{R})(\forall z \in \mathbb{R})x(y + z) = xy + xz$.

FA7. $(\forall x \in \mathbb{R})(\forall y \in \mathbb{R})(x - y) + y = x$.

FA8. $(\forall x \in \mathbb{R})(\forall y \in \mathbb{R})(y \neq 0 \Longrightarrow (x \div y).y = x)$.

FA9. $(\forall x \in \mathbb{R})x + 0 = x$.

FA10. $(\forall x \in \mathbb{R})x.1 = x$.

FA11. $0 \neq 1$.

4 A second detour into Logic: five logical rules for proofs

We now start presenting the rules of Logic that govern proofs. Eventually, when we have gone through the full list, it will turn out that **there are exactly 15 rules**, all of which are very easy to remember and understand.

We have already seen one logical rule (Rule SEE). We now present five more. So when you have finished studying this section you will know six logical rules. The remaining nine logical rules will be presented in later lectures, but **you do not need to worry that new rules, dozens of rules, will keep appearing every time.** There are only 15 logical rules, and we will have discussed all of them pretty soon, and once we are done with them there will be no more logical rules.

4.1 Proving universal sentences

We are about to present a rule for proving universal sentences. But, before we explain the rule, let us say a few words about naming sentences.

Using letters to name sentences

If we want to talk repeatedly about an object, we give it a short name, usually a letter. (We have been doing this already, dozens of times. For example, we have been talking a lot about real numbers, and we used for them such names as "x", "y", and "z".) This applies to sentences as well as to numbers. For example, we can take the sentence "x is positive" (i.e., "$x > 0$") and call it P. But it is more common to give a sentence a name that tells you which variables occur in it. So, instead of giving the sentence "$x > 0$" the name P, we call it $P(x)$. (We do this if we are interested in discussing the role of x. If we are not, then we can just call the sentence P.) Similarly, we could give the sentence "$x < 1$" the name "$Q(x)$", and the sentence "$x^2 < x$ the name "$R(x)$". Then the sentence "$(\forall x \in \mathbb{R})((P(x) \wedge Q(x)) \Longrightarrow R(x))$" is the sentence "$(\forall x \in \mathbb{R})((x > 0 \wedge x < 1) \Longrightarrow x^2 < x)$", that is "if x is a real number then if x is positive and less than one then $x^2 < x$" or, even better: "if x is a real number which is positive and less than one then $x^2 < x$".

And now we can state the rule:

The rule for proving a universal sentence (Rule $\forall_{prove}$, also known as the "universal generalization" rule)

If $P(x)$ is a sentence involving the variable x, then:

1. If, starting with "Let x be arbitrary" you prove $P(x)$, then you can conclude that

$$(\forall x)P(x)\,.$$

2. If S is a set, and starting with "Let x be an arbitrary member of S" you prove $P(x)$, then you can conclude that $(\forall x \in S)P(x)$.

(The meaning of "arbitrary" is explained in the next box.)

Example 7. To prove that $(\forall x \in \mathbb{R})x^2 \geq 0$ you start with

Let $x \in \mathbb{R}$ be arbitrary,

or, even better,

Let be an arbitrary real number,

and work your way to the sentence

$x \geq 0$.

If you manage to do that, then you can go to the conclusion that

$(\forall x \in \mathbb{R})x^2 \geq 0\,.$ □

What are "arbitrary" objects?

In order to prove that a property P is true for every object in some set S, we pick an "arbitrary" member of S, call it x, and prove that P holds for x. (You could call it any name you want: x, or y, or a, or α, or "Billy". The name doesn't matter, except for one thing: you *cannot* use as a name a symbol that is already the name of something else.) If you manage to do this, then you may conclude that P is true for every member of S. This is called the *universal generalization rule*, and will be widely used throughout this course, because it is one of the most important logical rules.

An "arbitrary" member of S a member of S that we can work with and reason about, but we don't know which specific object it is, and for all we know could be any member of S. You can think of this as follows: an "arbitrary" member of S is a member of S that has been given to you by an imaginary character called the CAT ("creator of arbitrary things"), who brings this object over to you inside a sealed envelope, so you have the object in your hands and can reason about it, but you don't know which member of S it is, and could turn out to be any member of S. Therefore, whatever you say about this object had better be true of *every* member of S, because if there is just one member of S for which what you say isn't true, then the "arbitrary" object could turn out to be that object.

Another way to think about "arbitrary" objects is this: imagine that x is a member of S that is going to be brought to you by the CAT *later*, after you have written your proof. So when you write your proof whatever you say about x had better be true for *all* members of S, because if there is one member of S for which what you say isn't true (such a member of S is called a "counterexample") then that member of S could be precisely the one that the CAT gives you.

You can even go farther, and think that the CAT is very mean, and wants to prove you wrong. So the CAT will look for a counterexample and will give you that counterexample. The only way you can outsmart the CAT is by making sure that what you say is true *for all members of* S, so that the CAT cannot find a counterexample.

Arbitrary sentences

Naturally, there can be arbitrary sentences as well as arbitrary real numbers, arbitrary sets, arbitrary functions, arbitrary cows, ... , arbitrary anything.
For example, if we say "Let $P(x)$ be an arbitrary sentence", we mean that $P(x)$ is a sentence that could be any sentence, and we don't know which sentence it is.

IMPORTANT CONVENTION. When we name a sentence $P(x)$, then, if a is the name of an object, we use $P(a)$ for the sentence obtained by "plugging in" a for x in $P(x)$.

Example 8. If we give the name $P(x)$ to the sentence "$x^2 < x$", then $P(3)$ stands for the sentence "$3^2 < 3$" (which, of course, is false) and $P(\frac{1}{2})$ stands for the sentence "$\left(\frac{1}{2}\right)^2 < \frac{1}{2}$" (which is true). And, if a is some real number (that could be a specific real number or an arbitrary one), then "$P(a)$" stands for the sentence "$a^2 < a$".

4.2 Proving a conjunction: a stupid but important rule

The rule for proving a conjunction (Rule $\wedge_{prove}$)

If P, Q are sentences, and you have proved P and you have proved Q, then you are allowed to go to $P \wedge Q$.

IMPORTANT REMARK. You may wonder "what is the point of such a rule?" But you cannot dispute that it is a reasonable rule! Of course, if you know that "today is Friday" and you also know that "tomorrow is Saturday", then you will have no doubt that "today is Friday and tomorrow is Saturday" is true. So you should have no problem accepting (and remembering) this rule. You may not understand why it is needed. So let me tell you why. Suppose it was a computer doing proofs, rather than a human being like you. Suppose the computer is told that today is Friday and then it is told that tomorrow is Saturday. How will the computer know that it can write "today is Friday and tomorrow is Saturday". It won't, unless you tell it. Computers do not "know" anything on their own. If you want the computer to "know" that once it knows that "today is Friday" and also that "tomorrow is Saturday", then it can write "today is Friday and tomorrow is Saturday", then you have to **tell** the computer. In other words, you have to input Rule $\wedge_{prove}$ into the computer. Proofs are mechanical manipulations of strings of symbols, and should therefore be doable by a computer. So Rule $\wedge_{prove}$ is needed.

And now let's go back to you, the human being. How do *you* know that, once you find out that "today is Friday" and also that "tomorrow is Saturday", then you can say (or write) "today is Friday and tomorrow is Saturday". **You know this because you know Rule** $\wedge_{prove}$. You know this rule so well, it is embedded so deeply in your mind, that you don't even realize that the rule is there. But **the rule is there!**

Here is another way to think about this. Suppose you didn't know any English at all. Then you would not know what the word "and" means, and you would not know that, if you have two sentences P and Q, then you can say or write "P and Q". As you learn English, at some point you would learn the meaning of the word "and" and then you would learn that when you have two sentences P and Q, then you can say or write "P and Q". (And I would even argue that this rule about that use of "and" is in fact what "and" means, but I will not pursue this now.) The point is: *there are* rules for using the word "and", and those rules have to be *learned*, and they only look obvious to you because you already learned them a long time ago and have grown accustomed to them.

What we are doing in Logic is **elucidating the laws of thought, making them explicit, bringing them to the surface, as it were**, so that we can, for example, pass them on from our minds to a computer: the computer does not "know" any of the things that you know, unless you tell the computer those things. And this applies even to the rules that you know so well that they are deeply embedded in your subconscious, so you take them for granted without even realizing that there is something to be known there.

Once you understand this, you will also see that **it is not an accident that modern Logic developed first, at the end of the 19th century and the beginning of the 20th century, and computers came into being soon afterwards.** □

4.3 Proving an implication

The rule for proving an implication
(Rule $\Longrightarrow_{prove}$)

Suppose P, Q are sentences. Suppose you start a proof with "Assume P", and you prove Q. Then you can go to $P \Longrightarrow Q$.

Example 9. Say you are a Martian who just landed on Earth, you know nothing about the days of the week, and you want to prove that to your own satisfaction that "If today is Friday then tomorrow is Saturday". To apply Rule $\Longrightarrow_{prove}$, you would begin by "assuming that today is Friday." This means that you would imagine that today is Friday, and see what would happen in that case. For example, you could go to a public library and look at lots of newspapers published on a Friday, and you would see that every time such a paper talks about the following day it says something like "tomorrow is Saturday." Then you would be reasonably confident that the sentence "If today is Friday then tomorrow is Saturday" is true. And it would not matter whether today is Friday or not.$\square$

For a more mathematical example of a proof using Rule $\Longrightarrow_{prove}$, see the proof of Theorem 4 on page 30, where we show you an example that uses both Rules $\Longrightarrow_{prove}$ and $\vee_{prove}$.

4.4 Proving a disjunction

The rule for proving a disjunction (Rule $\vee_{prove}$)

Suppose P and Q are sentences, and you want to prove $P \vee Q$. Here is what you can do. You look at the two possible cases, when P is true and when P is false. If P is true then of course $P \vee Q$ is true, so we are O.K. So all we have to do is look at the other case, when P is false, and prove that in that case Q is true.
So here is the rule:

I. If, assuming that P is false, you can prove Q, then you can go to $P \vee Q$.

II. If, assuming that Q is false, you can prove P, then you can go to $P \vee Q$.

4.5 Proofs by contradiction

One way to prove a statement is to try to imagine a world in which the statement isn't true, and show that that world is impossible. This is called **proof by contradiction**

An example of such a proof was given in Lecture 1, in the proof of Euclid's theorem that the set of prime numbers is infinite. The way we proved it was by assuming (i.e., imagining) that the set was finite, and showing that a world in which this happens is an impossible world, because something impossible has to happen in it. (In our case, the impossible thing that had to happen in that world was that there is a list L for which the following two things are true:

(I) L is a list of all the primes,

and

(II) there is a prime that is not on the list L.

Clearly, a world where both (I) and (II) happen cannot exist.) It follows that the statement that the set of primes is finite cannot be true, so the set of primes is infinite.

Here is the precise statement of the rule.

THE PROOF BY CONTRADICTION RULE

(I) Suppose that P is a sentence that you want to prove. Suppose you can do the following: Assume $\sim P$ (that is, assume that P is false) and prove another statement Q as well as its negation $\sim Q$. Then you can include P as a step in your proof.

(II) Suppose that P is a sentence, and you want to prove that P is false, that is, that $\sim P$. Suppose you can do the following: Assume P (that is, assume that P is true) and prove another statement Q as well as its negation $\sim Q$. Then you can include $\sim P$ as a step in your proof.

The structure of a proof by contradiction is then as follows:

,

Assume $\sim P$.
$\vdots$
(Insert a proof of Q using $\sim P$ here.)
$\vdots$
Q
$\vdots$
(Insert a proof of $\sim Q$ using $\sim P$ here.)
$\vdots$
$\sim Q$

P

or as follows

,

Assume P.

$\vdots$

(Insert a proof of Q using $\sim P$ here.)

$\vdots$

Q

$\vdots$

(Insert a proof of $\sim Q$ using $\sim P$ here.)

$\vdots$

$\sim Q$

$\sim P$

5 Using the field axioms to define new things and prove new theorems

In the study of real numbers, all the things we want to talk about that are not primitive concepts have to be introduced into the theory by defining them, i.e., by explaining what they mean.

One concept that does not appear in the list of basic concepts is the *negative*, or *additive inverse* of a real number. The "$-$" symbol does appear, but only in the context of the subtraction operation, which takes two numbers x, y and produces the number $x = y$. The "negative of" concept, as when we talk about the number $-x$ for a given number x, is different. More precisely: subtraction is a *binary operation*. (Binary operations are discussed in page 1.) But "minus", or "negative of", is a one-argument operation. So they are different. The binary operation "$-$" is the one that appears in the axioms. The other one, the one-argument operation "$-$", does not, so we have to introduce it properly by defining what it means.

But do not worry. Defining $-x$ is going to be simple! All we will do is say that "$-x$" means "$0 - x$". Easy, isn't it?

A similar situation arises with the "multiplicative inverse" concept. The field axioms do not mention x^{-1}, so we have to introduce this concept by explaining what it means, i.e., by *defining* it. And, again, it is going to be easy. We will just say that "x^{-1}" means "$1 \div x$". But now there is a detail we have to take care of. The additive inverse $-x$ exists for every real number x. But the multiplicative inverse x^{-1} does not exist if $x = 0$. So in the definition of x^{-1} we have to specify that $x \neq 0$.

5.1 Definition of "$-x$" and "x^{-1}"

Definition 2. If x is a real number, then the negative of x is the number $-x$ given by

$$-x = 0 - x\,. \qquad \square$$

Notice that this defines $-x$ in terms of the basic concepts, because it involves "0" and "−" (the sense of difference of two numbers, not that of the negative of a number).

And we can do the same for the "inverse" of a nonzero real number:

Definition 3. If x is a real number, and $x \neq 0$, then the inverse (also called multiplicative inverse) of x is the number x^{-1} given by

$$x^{-1} = 1 \div x\,.$$

Again, this defines x^{-1} in terms of the basic concepts, because it involves "1" and division. □

5.2 The cancellation laws for addition and multiplication

Warning. You are going to find the next few profs extremely silly. For example, we are going to prove that 2+2 = 4, and you will probably complain, saying: "That's silly! I already know that, so what's the point of proving it?" I have three answers to that. First answer: If it's not in the axioms then we have to prove it. Second answer: Think of this course as similar to a language course, in which you are learning a new language. In, for example, a course of English as a foreign language, you would not start the course with Shakespeare. You would start with simple sentences like "the cat is on the mat", "my mom loves me", or "Jack and Jill went up the hill." Then, step by step, you would move on to harder, more complicated sentences, and maybe by the end of the semester you would be reading Hamlet's soliloquy. In this course, we are doing the same thing. Proving that 2+2=4 is the equivalent of learning to read and write the statement "the cat is on the mat." And, believe me, by the end of the semester we will be doing really sophisticated proofs. □

Our first two theorems are going to be about *cancellation*. The first theorem says that if you have an equality

$$x + y = x + z$$

then you can "cancel" the x and go to

$$y = z\,.$$

The second theorem says the same thing about multiplication: if you have an equality

$$xy = xz$$

then you can "cancel" the x and go to

$$y = z\,,$$

provided that $x \neq 0$. This qualification is important: if $x = 0$ then in general you cannot cancel the x. For example, it is true that $0 \times 3 = 0 \times 5$ (because $0 \times 3 = 0$ and $0 \times 5 = 0$, so $0 \times 3 = 0 \times 5$). But you cannot cancel the zero and conclude that $3 = 5$.

The proofs of these two theorems are very similar. So I will give you the first proof, and ask you to do the second one. The proof of the second theorem is almost exactly the same as that of the first theorem, but you have to deal with the important difference between the two: in the second proof, you will need the condition that $x \neq 0$.

Theorem 1. *(The cancellation law for addition.) If x, y, z are real numbers, and $x + y = x + z$, then $y = z$. (In formal language: this says that* $(\forall x \in \mathbb{R})(\forall y \in \mathbb{R})(\forall z \in \mathbb{R})(x + y = x + z \Longrightarrow y = z)$.*)* □

The idea of the proof. It's very simple. You take the equality $x+y = x+z$, add $-x$ to both sides, and you get $(-x) + x + y = (-x) + x + z$. Then $-x$ and x cancel, and you get $y = z$. **That's all there is to it!**

And now we have to do the same thing making sure that we use the axioms, the logical rules, the definitions we have given so far, and nothing else. (In general, for a proof, we would also be allowed to use the theorems we have proved before, but at this point we are proving our first theorem, so there aren't any "theorems proved before".)

PROOF. Let x, y, z be arbitrary real numbers.
Assume that $x+y=x+z$. We want to prove that $y=z$.

We have

$$(5.4)\qquad (-x)+(x+y)=(-x)+(x+y)\,,$$

because of Axiom EA1 (every object is equal to itself), applied to the object $(-x)+(x+y)$.

Since $x+y=x+z$, we can substitute $x+z$ for the second of the two $x+y$'s in (5.4), and conclude that

$$(5.5)\qquad (-x)+(x+y)=(-x)+(x+z)\,.$$

The associative law of addition implies that

$$(-x)+(x+y)=((-x)+x)+y\,,$$

so we may substitute $((-x)+x)+y$ for $(-x)+(x+y)$ in (5.5), and conclude that

$$(5.6)\qquad ((-x)+x)+y=(-x)+(x+z)\,.$$

Similarly,

$$(-x)+(x+z)=((-x)+x)+z\,,$$

so

$$(5.7)\qquad ((-x)+x)+y=((-x)+x)+z\,.$$

Now, according to the definition of "negative", we have $-x=0-x$. So we may substitute $0-x$ for $-x$ in (5.7), and get

$$(5.8)\qquad ((0-x)+x)+y=((0-x)+x)+z\,.$$

Next, according to Axiom FA7 (applied with x in the role of y, and 0 in the role of x), we have

$$(0-x)+x=0\,.$$

So we may substitute 0 for $(0-x)+x$ in (5.8), and get

$$(5.9)\qquad 0+y=0+z\,.$$

Finally, according to Axiom FA9[7] $0+y=y$ and $0+z=z$. So we may substitute y for $y+0$ and z for $z+0$ in (5.9), getting

$$(5.10)\qquad y=z\,.$$

[7] To be precise, Axiom FA9 tells us that $y+0=y$ and $z+0=z$. To get the conclusions that $0+y=y$ and $0+z=z$. we need a couple of extra steps, using Axiom FA3 to conclude that $y+0=0+y$ and $z+0=0+z$, so , $0+y=y$ and $0+z=z$. **What we have done in this proof is something that we will keep doing from now on: skip steps that are trivial and obvious.**

So we have proved that $y = z$, assuming that $x + y = x + z$. This means, by Rule $\Longrightarrow_{prove}$, that we have proved

$$x + y = x + z \Longrightarrow y = z\,.$$

And this has been proved for arbitrary real numbers x, y, z. So Rule $\forall_{prove}$ allows us to conclude that

$$(\forall x \in \mathbb{R})(\forall y \in \mathbb{R})(\forall z \in \mathbb{R})(x + y = x + z \Longrightarrow y = z)\,.$$

Q.E.D.

THE SAME PROOF, WRITTEN MORE CONCISELY, SKIPPING LOTS OF TRIVIAL STEPS. Assume that $x + y = x + z$. We want to prove that $y = z$.
We have (thanks to Axiom EA1, with the object $(-x) + (x + y)$ in the role of x):

$$(-x) + (x + y) = (-x) + (x + y)\,. \tag{5.11}$$

Since $x + y = x + z$, we get

$$(-x) + (x + y) = (-x) + (x + z)\,. \tag{5.12}$$

Using the associative law of addition, we find

$$((-x) + x) + y = (-x) + (x + y)\,, \tag{5.13}$$

and

$$((-x) + x) + z = (-x) + (x + z)\,, \tag{5.14}$$

hence Rule SEE allows us to write

$$((-x) + x) + y = ((-x) + x) + z\,. \tag{5.15}$$

Since $-x = 0 - x$, (5.13) implies

$$((0 - x) + x) + y = ((0 - x) + x) + z\,. \tag{5.16}$$

According to Axiom FA7, $(0 - x) + x = 0$. So

$$0 + y = 0 + z\,. \tag{5.17}$$

Since $0 + y = y$ and $0 + z = z$, we get

$$y = z\,. \tag{5.18}$$

Q.E.D.

Theorem 2. *(The cancellation law for multiplication.) If x, y, z are real numbers, $x \neq 0$, and $xy = xz$, then $y = z$.* □

PROOF. YOU DO THIS ONE. (It's almost exactly the same as the previous proof.)

Problem 1. *Prove Theorem 2.* □

5.3 Some simple but tricky proofs

Theorem 3. *If $x \in \mathbb{R}$, then $x.0 = 0$. (In formal language, this says that $(\forall x \in \mathbb{R})x.0 = 0$.)* □

Remark. This proof is short and easy, **but it involves a trick**. So **you have to know the trick**, because if you are asked to write this proof in an exam[8] and you don't know the trick, you may not be able to figure it out on your own. This means that **you have to study this proof**[9]**.**

PROOF. We are going to use Rule $\forall_{prove}$. In order to prove a sentence of the form "$(\forall x \in \mathbb{R})$blahblahblah" we start by letting x be an arbitrary real number, and prove "blahblahblah".

Let $x \in \mathbb{R}$ be arbitrary.

We apply Axiom EA1 to write

$$x.0 = x.0\,. \tag{5.19}$$

Then we use Axiom FA9 (with 0 in the role of x), to conclude[10] that

$$0 + 0 = 0\,. \tag{5.20}$$

[8]Which may very well happen, believe me! I *do* know what I am talking about!

[9]Actually, **you should study all the proofs**, but this particular one is tricky, so you have to study it very carefully.

[10]**This is the trick I told you about!**. If you don't know the trick you probably will not be able to figure out the proof. If you ask "why do you do this?, why do you write $0 + 0 = 0$?", the answer is "because it works". How did I figure out that this is what one has to do? **I didn't.** Some very smart mathematician figured it out first, and taught it to that mathematician's students, and the knowledge was passed on until I learned this trick in an undergarduate Algebra course many years ago.

Then we use Rule SEE to substitute $0+0$ for 0 in the left side of (5.19), getting

$$x.(0+0) = x.0\,. \tag{5.21}$$

Next we use the distributive law (Axiom FA6) to conclude that

$$x.(0+0) = x.0 + x.0\,. \tag{5.22}$$

Then, using Rule SEE again, we find

$$x.0 + x.0 = x.0\,. \tag{5.23}$$

But Axiom FA9 implies that

$$x.0 + 0 = x.0\,. \tag{5.24}$$

Hence, using Rule SEE, we obtain

$$x.0 + x.0 = x.0 + 0\,. \tag{5.25}$$

We now use the cancellation law of addition (Theorem 1), with $x.0$ in the role of x, $x.0$ in the role of y, and 0 in the role of z, to conclude that

$$x.0 = 0\,. \tag{5.26}$$

So we have proved that $x.0 = 0$ assuming that x was an arbitrary real number. Hence by Rule $\forall_{prove}$, we can conclude that

$$(\forall x \in \mathbb{R})x.0 = 0\,.$$

Q.E.D.

Theorem 4. *If $x \in \mathbb{R}$, $y \in \mathbb{R}$, and $xy = 0$, then $x = 0$ or $y = 0$.* □

COMMENT: In formal language, we want to prove

$$(\forall x \in \mathbb{R})(\forall y \in \mathbb{R})(xy = 0 \Longrightarrow (x = 0 \vee y = 0))\,. \tag{5.27}$$

Here is how we can prove this.

First, since we want to prove a universal sentence, we introduce arbitrary real numbers x and y, and set out to prove that

$$xy = 0 \Longrightarrow (x = 0 \vee y = 0) \,. \tag{5.28}$$

Since (5.28) is an implication, we will want to use the rule for proving an implication, that is, Rule $\Longrightarrow_{prove}$. This means that we will assume that $xy = 0$, and try to prove that $x = 0 \vee y - 0$.

Next, in order to prove the disjunction $x = 0 \vee y = 0$, we will use Rule $\vee_{prove}$: we will assume that $\sim x = 0$ (that is, that $x \neq 0$) and try to prove that $y = 0$.

And here is the proof.

PROOF:
Let x, y be arbitrary real numbers.
Assume that $xy = 0$.
We want to prove that $x = 0 \vee y = 0$.
Assume that $\sim x = 0$, i.e., that $x \neq 0$.
We want to prove that $y = 0$, so as to apply Rule $\vee_{prove}$.
Since $x \neq 0$, we have $1 \div x \in \mathbb{R}$, by Axiom FA1.d.
Also, we know that $xy = 0$.
Using the fact that any real number times zero equals zero, we get

$$(1 \div x).0 = 0\,. \tag{5.29}$$

Since $xy = 0$, we can use Rule SEE and conclude that

$$(1 \div x).(xy) = 0\,.$$

From Axiom FA4, it follows that

$$((1 \div x).x)y = (1 \div x).(xy)\,.$$

Hence, using Rule SEE, we get $((1 \div x).x)y = 0$.
By Axiom FA8, we know that $(1 \div x)x = 1$.
So, using Rule SEE, we get $1.y = 0$.
But Axiom FA3 implies that $1.y = y.1$).
So, using Rule SEE again, we get $y.1 = 0$.
But Axiom FA10 implies that $y.1 = y$.
And then, using Rule SEE once more, we get $y = 0$.
Since we have proved that $y = 0$ assuming that $\sim x = 0$, Rule $\vee_{prove}$ allows us to conclude that $x = 0 \vee y = 0$.
Since we have proved that $x = 0 \vee y = 0$ assuming that $x \neq 0$, Rule $\Longrightarrow_{prove}$ allows us to conclude that

$$x \neq 0 \Longrightarrow (x = 0 \vee y = 0).$$

Finally, since we have proved that $x \neq 0 \Longrightarrow (x = 0 \vee y = 0)$ for arbitrary x, y, we can conclude that (5.27) is true.

Q.E.D.

5.4 The numbers 2,3,4, 5, and 6

The axioms do not mention any numbers other than 0 and 1. So, if we want to talk about the number 2, we have to introduce it first, by giving a definition.

And, again, this is very easy.

Definition 4. $2 = 1 + 1$. □

Definition 5. $3 = 2 + 1$. □

Definition 6. $4 = 3 + 1$. □

Theorem 5. $2 + 2 = 4$. □

PROOF. It follows from Axiom EA1 that

$$2 + 2 = 2 + 2 \,. \tag{5.30}$$

Definition 4 tells us that $2 = 1 + 1$.

So (using Rule SEE) we may substitute $1 + 1$ for the last of the four 2's of Equation (5.30), and get

$$2 + 2 = 2 + (1 + 1) \,. \tag{5.31}$$

By the associative law of addition (Axiom FA2), $2 + (1 + 1) = (2 + 1) + 1$. So (using Rule SEE)

$$2 + 2 = (2 + 1) + 1 \,. \tag{5.32}$$

But $2 + 1 = 3$, by Definition 5. Hence

$$2 + 2 = 3 + 1 \,. \tag{5.33}$$

And $3 + 1 = 4$, by Definition 6. Therefore

$$2 + 2 = 4 \,. \tag{5.34}$$

Q.E.D.

Theorem 6. $2 \times 2 = 4$. □

PROOF. We know from Theorem 5 that

(5.35) $$2 + 2 = 4\,.$$

Axiom FA9 tells us that

(5.36) $$2 \times 1 = 2\,.$$

Using Rule SEE, we can substitute 2×1 for 2 in (5.35), getting:

(5.37) $$2 \times 1 + 2 \times 1 = 4\,.$$

The distributive law (Axiom FA6) implies that

(5.38) $$2 \times (1 + 1) = 2 \times 1 + 2 \times 1\,.$$

So we may substitute $2 \times (1 + 1)$ for $2 \times 1 + 2 \times 1$ in (5.37), getting

(5.39) $$2 \times (1 + 1) = 4\,.$$

Definition 4 tells us that $2 = 1 + 1$.

So (using Rule SEE) we may substitute 2 for $1 + 1$ in Equation (5.39), and get

(5.40) $$2 \times 2 = 4\,.$$

Q.E.D.

Definition 7. $5 = 4 + 1$. □

Definition 8. $6 = 5 + 1$. □

Definition 9. $6 = 5 + 1$. □

Theorem 7. $3 + 3 = 6$.

PROOF. YOU DO THIS ONE.

Problem 2. *Prove Theorem 7.* □

Theorem 8. $3 \times 2 = 6$. □

PROOF. YOU DO THIS ONE.

Problem 3. *Prove Theorem 8.* □

6 The basic facts of real number theory. Part II: the order axioms

In Part I, in Section 1.3, we studied the properties of the real numbers that have to do with the four basic arithmetic operations ($+$, $\times$, $-$ and $\div$).

Now we must talk about the remaining symbol in the basic vocabulary of real number theory, namely, the *order relation* $<$ ("less than").

I will give the axioms governing $<$, and then we will prove some theorems. And, since by now you know a lot about formal language, I will give you the axioms directly in formal language.

THE ORDER AXIOMS FOR THE REAL NUMBERS

OA1. $(\forall x \in \mathbb{R})(\forall y \in \mathbb{R})(x < y \vee y < x \vee x = y)$

OA2. $(\forall x \in \mathbb{R})(\forall y \in \mathbb{R})(x < y \Longrightarrow \sim (x = y \vee y < x))$

OA3. $(\forall x \in \mathbb{R})(\forall y \in \mathbb{R})(\forall z \in \mathbb{R})((x < y \wedge y < z) \Rightarrow x < z)$

OA4. $(\forall x \in \mathbb{R})(\forall y \in \mathbb{R})(\forall z \in \mathbb{R})(x < y \Longrightarrow x + z < y + z)$

OA5. $(\forall x \in \mathbb{R})(\forall y \in \mathbb{R})(\forall z \in \mathbb{R})((x < y \wedge 0 < z) \Longrightarrow x \cdot z < y \cdot z)$

6.1 A digression: combining more than two things

6.1.1 Sums, products, disjunctions, conjunctions of three things

In Axiom OA1 there appears the sentence "$x < y \vee y < x \vee x = y$". What does this mean? Let me explain.

According to our previous discussion of the connectives "$\vee$" and "$\wedge$", in subsection 2.2, "$\vee$" and "$\wedge$" are *binary* operations[11] on the set of all statements: In exactly the same way as "$+$" and "$\times$" can be used to combine two real numbers x and y and produce the real numbers $x + y$, $x \times y$ (also written as $x.y$ or xy), the connectives "$\vee$" and "$\wedge$" are used to combine two sentences P, Q and produce the sentences $P \vee Q$ and $P \wedge Q$.

For addition and multiplication, one would like to add or multiply three or more numbers. But, strictly speaking, one can only add and multiply numbers by doing it two at a time, so if you want to add or multiply three numbers x, y, z, you cannot just write $x + y + z$ or xyz. You would have

[11] Binary operations are discussed on page 1.

to write $(x + y) + z$, or maybe $x + (y + z)$, and $(xy)z$, or maybe $x(yz)$. However, the associative laws for addition and multiplication tell us that $(x+y)+z = x+(y+z)$ and $(xy)z = x(yz)$. This makes it possible to *define* $x+y+z$ to mean $(x+y)+z$, or $x+(y+z)$, and xyz to mean $(xy)z$ or $x(yz)$. In both cases, it does not matter in which way you associate the numbers (first you add x and y, and then add z to the result, or first you add y and z, and then add x to the result, and similarly for multiplication).

So you can write:

Definition 10. If x, y, z are real numbers, then $x + y + z$ is the number $(x + y) + z$. □

Remark 5. It is easy to prove (and I will skip the proof) that if x, y, z are real numbers, then $x + y + z = x + (y + z)$. (The reason for this is that, according to the definition, $x + y + z = (x + y) + z$. And according to the associative law $(x + y) + z = x + (y + z)$, so $x + y + z = (x + y) + z$.) □

Definition 11. If x, y, z are real numbers, then xyz is the number $(xy)z$. □

Remark 6. It is easy to prove (and I will skip the proof) that if x, y, z are real numbers, then $xyz = x(yz)$. (The reason for this is, once again, the associative law.) □

In a similar way, since "$\vee$" and "$\wedge$" can in principle only be used to connect *two* sentences, it is not clear what a sentence such as "$x < y \vee y < x \vee x = y$" means.

But we can make it meaningful by defining what it means.

Definition 12. If P, Q, R are sentences, then "$P \vee Q \vee R$" is the sentence "$P \vee (Q \vee R)$". □

Definition 13. If P, Q, R are sentences, then "$P \wedge Q \wedge R$" is the sentence "$P \wedge (Q \wedge R)$". □

It turns out that in this case it is also true that you can "associate" differently. In some sense, "$P \vee (Q \vee R)$" is "the same" as "$(P \vee Q) \vee R$", and "$P \wedge (Q \wedge R)$" is "the same" as "$(P \wedge Q) \wedge R$".

But what exactly do we mean by "the same"? Certainly, the sentences "$x < y \vee (y < x \vee x = y)$" and "$(x < y \vee y < x) \vee x = y$" are not exactly the same. (For example, "$x < y \vee (y < x \vee x = y)$" starts with an "$x$", while "$(x < y \vee y < x) \vee x = y$" starts with a left parenthesis.)

It turns out that these sentences are "the same" in a way which is all we need for proof-writing:

- **Any time one of the sentences "$P \vee (Q \vee R)$", "$(P \vee Q) \vee R$" is true, it follows that the other one is true as well, so when you are writing a proof and get to one of those sentences, you can pass to the other one.**
- **Any time one of the sentences "$P \wedge (Q \wedge R)$", "$(P \wedge Q) \wedge R$" is true, it follows that the other one is true as well, so when you are writing a proof and get to one of those sentences, you can pass to the other one.**

The precise meaning of "the same" involved in this is called "logical equivalence". This concept will be explained later, but what you need to know now is that **when two sentences are logically equivalent you can move freely from one to the other**.

6.1.2 Combining more than three things

And now, as I am sure you must have guessed, we could move on and define addition and multiplication of four real numbers. (We could define $x + y + z + w$ to mean, for example, $(x + y) + (z + w)$, or $x + (y + (z + w))$, or $((x + y) + z) + w$, and then prove that all these different expressions give rise to the same number, so in the end it does not matter which one you chose.)

And we could also define the product $xyzw$ of four numbers.

And we could define the disjunction $P \vee Q \vee R \vee S$ of four sentences P, Q, R, S, or the conjunction $P \wedge Q \wedge R \wedge S$.

And then we could go on and on and on, to the sum and product of five numbers, and the conjunction and disjunction of five sentences.

And then move on to six numbers and six sentences.

And on and on and on.

But at some point we will need a better way!

- We would like to be able to talk about the sum and the product of n real numbers, where n is an arbitrary[12] natural number.

[12]By now, you know what "arbitrary" means, don't you?

- And we would like to talk about the disjunction and the conjunction of n sentences, where n is an arbitrary natural number.

Can we do this? **Yes, we can!** But we need to find a language to talk about an arbitrary number of numbers, or an arbitrary number of sentences. We are going to do it in this course, in few lectures from now.

6.1.3 The trouble with other operations: the need for parentheses

Can we do for subtraction and division what we did in the previous section for addition and multiplication. For example, can we "subtract three numbers"? If you haven't thought about this issue until now, think about it now!

If I tell you "add the numbers 3, 6, and 2", you know what to do, and you will see right away that the answer is 11. But if I ask you "subtract the numbers 3, 6, and 2", you probably don't know what to do, because you don't even understand the question. **There is no natural way to assign a meaning to the difference of three numbers x, y, z.** You can try $(x - y) - z$, or $x - (y - z)$, for example, but these two are different. (For example: $(3 - 6) - 2 = -5$, but $3 - (6 - 2) = -1$.)

And the ultimate reason for this is that **the operations of subtraction and multiplication are not associative.** And this is why you cannot do for them what we did for addition and multiplication.

And a similar problem arises with $\Longrightarrow$, which is a binary operation in the set of sentences. **Implication is not an associative operation: If P, Q, R are three sentences, then $P \Longrightarrow (Q \Longrightarrow R)$ is, in general, not logically equivalent to $(P \Longrightarrow Q) \Longrightarrow R$.**

Example 10. If P, Q, R are sentences, and P happens to be false, while R is true (and it doesn't matter what Q is), then $P \Longrightarrow (Q \Longrightarrow R)$ is true, and $(P \Longrightarrow Q) \Longrightarrow R)$ is false[13]. So the two sentences are not logically equivalent. □

[13] Why is $P \Longrightarrow (Q \Longrightarrow R)$ true? Wny is $P \Longrightarrow (Q \Longrightarrow R)$ false? We are going to learn to deal with these questions when we do truth tables later, but here is a quick way to see it: $A \Longrightarrow B$ is true if B is true, irrespective of whether A is true or false, and $A \Longrightarrow B$ is true if A is false, irrespective of whether B is true or false. And the only case when $A \Longrightarrow B$ is false in when A is true and B is false. Applying these rules, we can easily see that, if P is false then $P \Longrightarrow (Q \Longrightarrow R)$ is true. And if, in addition, R is false, then $P \Longrightarrow Q$ is true, so $(P \Longrightarrow Q) \Longrightarrow R$ is false.

And **this is the main reason why we need parentheses**: to distinguish between $(x-y)-z$ and $x-(y-z)$, and between $(P \Longrightarrow Q) \Longrightarrow R$ and $P \Longrightarrow (Q \Longrightarrow R)$.

6.2 Explanation of the order axioms

We now go back to the order axioms for $\mathbb{R}$, and explain the meaning of these axioms.

Axiom OA1 says that any two real numbers can always be compared by means of the binary relation "$<$":

> If we are given two real numbers x, y, one of the three possibilities
>
> $$x < y\,, \qquad y < x\,, \qquad x = y$$
>
> must occur.

Let us contrast this with other binary relations that do not have this property. For example, let us look at the binary relation "divides", on the set $\mathbb{Z}$. (Recall that, by definition, "x divides y" means "there exists an integer z such that $y = xz$".) For this relation, it is *not* true that any two integer can be compared. For example, if you take $x = 5$ and $y = 7$, it is not true that one of the three possibilities "x divides y", "y divides x", "$x = y$", holds. (Indeed, 5 does not divide 7, 7 does not divide 5, and 5 and 7 are not equal.)

Axiom OA2 says that

> If we are given two real numbers x, y, only one of the three possibilities
>
> $$x < y\,, \qquad y < x\,, \qquad x = y$$
>
> can occur.

Why do I say that the axiom says that? The axiom, in principle, looks quite different. What it says is that, if you are given two real numbers x, y, then

$$x < y \Longrightarrow \sim (x = y \vee y < x)\,.$$

That is: "if $x < y$ then it's not the case that $x = y$ or $y < x$." Using this, let us **prove** that only one of the three possibilities $x < y$, $y < x$, $x = y$ can occur.

PROPOSITION 1. If $x \in \mathbb{R}$ and $y \in \mathbb{R}$, only one of the three possibilities $x < y$, $y < x$, $x = y$ can occur.

Remark 7. I call this a "proposition" rather than a theorem because it is a very minor observation, unworthy of full-fledged theoremhood. □

PROOF. Suppose first that $x < y$. Then the Axiom OA2 tells us that neither "$x = y$" nor "$y < x$" can be true. So, in that case, only one of the three possibilities holds true.

Next, suppose that $y < x$. Then we are in exactly the same situation as in the previous paragraph, except that now y is in the role of x and x is in the role of y. In other words, the axiom tells us that $y < x \Longrightarrow \sim (y = x \vee x < y)$. Since we are assuming that $y < x$, it follows that it's not the case that $y = x$, or that $x < y$. So, once again, only one of the three possibilities $x < y$, $y < x$, $x = y$ can occur.

Next, suppose that $x = y$. In that case, the axiom tells us both that

$$x < y \Longrightarrow \sim (x = y \vee y < x) \tag{6.41}$$

and that

$$y < x \Longrightarrow \sim (y = x \vee x < y)\,. \tag{6.42}$$

(The second equation is the same as the first one, with x in the role of y and y in the role of x.) Since $x = y$, it cannot be the case that $x < y$, because if it was true that $x < y$ then (6.41) would imply that "$x = y$" is not true, contradicting the fact that we are assuming that $x = y$. Similarly, "$y < x$" cannot be true either, because if it was true then (6.42) would tell us that "$y = x$" is not true, but we are assuming that $x = y$, and then $y = x$.

So we have shown that if any one of the three possibilities occurs, then the other two do not occur. This completes our proof. **Q.E.D.**

Often, Axioms OA1 and OA2 are combined into a single statement:

AXIOMS OA1 AND OA2 COMBINED

If we are given two real numbers x, y, then one and only one of the three possibilities

$$x < y\,, \qquad y < x\,, \qquad x = y$$

occurs.
This statement is called the ***trichotomy property***. (Exactly as the word "dichotomy" applies when one and only one of two possibilities occurs, the word "trichotomy" tells us that one and only one of three possibilities occurs.)

Axiom OA3 says that the binary relation $<$ has the *transitive property.* (Another relation we already know that has the transitive property is equality: if $x = y$ and $y = z$ then $x = z$.)

Finally, **Axioms OA4 and OA5** tell us how $<$ interacts with the operations of addition and multiplication:

- **Axiom OA4** says you can add any real number z to both sides of an inequality $x < y$, and when you do that you get an inequality of the same kind, namely, $x + z < y + z$.

- **Axiom OA5** says almost the same thing about multiplication: you can multiply add both sides of an inequality $x < y$ by a real number z, and when you do that you get an inequality of the same kind, namely, $xz < yz$, *provided that z is positive*[14]

7 A third detour into Logic: four new logical rules

We now present four more logical rules. As we have already explained, when we have gone through the full list of logical rules, it will turn out that there are exactly 15 of them. (And they are all very easy to remember and understand.)

[14] There is a very good reason for that. If z is not positive then it is not true that when $x < y$ it follows that $xz < yz$. For example, if $x = 3$, $y = 4$, and $z = -1$, then $x < y$, but $xz = -3$ and $yz = -4$, so it is not true that $xz < yz$.

We have already seen six rules. So when you have finished studying this section you will know ten logical rules. The remaining five logical rules will be presented in later lectures.

We have already seen four "prove" rules. These are rules that can be used to prove sentences of a particular form. For example, to prove the sentence $(\forall x \in \mathbb{R})x.0 = 0$, we used—among other things—the rule for proving universal sentences, for which we are using the very easy to remember name "Rule $\forall_{prove}$". (According to this rule, if you start by writing "Let x be an arbitrary real number", and work your way to "$x.0 = 0$", then you are allowed to conclude that $(\forall x \in \mathbb{R})x.0 = 0$.) And there were three other "prove" rules: Rules $\wedge_{prove}$, $\vee_{prove}$, and $\Longrightarrow_{prove}$.

We will now present the "use" counterparts of these four "prove" rules. These are the rules that tell you what you can do if you have a sentence involving one of the four connectives $\forall$, $\wedge$, $\vee$, $\Longrightarrow$. Naturally, I am going to give these rules the names "Rule $\forall_{use}$", "Rule $\wedge_{use}$", "Rule $\vee_{use}$", and "Rule $\Longrightarrow_{use}$".

7.1 Using universal sentences

The rule for using a universal sentence (Rule $\forall_{use}$, a.k.a. the *specialization rule*)

If $P(x)$ is a sentence involving the variable x, t is a term, and $P(t)$ is the sentence obtained from $P(x)$ by substituting t for x, then

1. If you have, in an earlier step of your proof, the sentence $(\forall x)P(x)$, then you can go to $P(t)$.
2. If S is a set, and you have, in an earlier step of your proof, the sentence $(\forall x \in S)P(x)$, and if in addition you know that $t \in S$, then you can go to $P(t)$.

Example 11. Suppose you know that "$(\forall x)x = x$" (that is, every object is equal to itself), and you have a term that is the name of a particular object (say, the number 0). Then you can apply Rule $\forall_{use}$ and conclude that $0 = 0$. $\square$

Example 12. Suppose you know the (obviously true) fact that "All Rutgers professors are very smart". That is, if you use $S(x)$ to stand for "x is very smart". and use RP to stand for the set of all Rutgers professors, then you know that $(\forall x \in RP)S(x)$. Then Rule $\forall_{use}$ tells you that you can conclude that "H. J. Sussmann is very smart", that is, that $S(t)$, where "t" stands for "H. J. Sussmann".

The specialization rule says something very simple:

> If you know that something is true in general, that is, for all objects of a certain kind, then you can conclude that it is true in any special case, that is, for a particular object of that kind."

That is why Rule $\forall_{use}$ is called the **specialization rule.**

7.2 Using a conjunction: another stupid but important rule

> **The rule for using a conjunction (Rule $\wedge_{use}$)**
>
> If P, Q are sentences, and you have proved $P \wedge Q$, then you are allowed to go to P, and you are also allowed to go to Q.

IMPORTANT REMARK. This looks like a very stupid rule. But you should reread the "Important Remark" on Page 19, where we talked about another "stupid rule", namely, Rule $\wedge_{prove}$. That remark also applies to Rule $\wedge_{use}$. □

7.3 Using an implication

We now come to one of the most important rules in Logic: the rule for using an implication. For us, this rule will be called— guess what!—"Rule $\Longrightarrow_{use}$", but it also has a couple of much more impressive names: **Modus Ponens**, and **implication elimination**[15]

[15] "Modus Ponens" is an abbreviation of "modus ponendo ponens", which is Latin for "the way that affirms by affirming".

The rule for using an implication (Rule $\Longrightarrow_{use}$, a.k.a. *Modus Ponens*)

Suppose P, Q are sentences. Suppose you have the sentences $P \Longrightarrow Q$" and "P" in previous steps of your proof. Then you can go to Q.

Example 13. Suppose you know that "If you are a student then you are entitled to a discount" and you also know that you are a student. Then you can conclude that you are entitled to a discount.

7.4 The "for all...implies" one-two punch

One of the most important and widely used combinations of moves in proofs us what we may call ***the "for all...implies" one-two punch***.

It works like this:

- First, you bring into your proof a statement S of the form "for every x of some kind, if something happens then something else happens". That is, $(\forall x)(A(x) \Longrightarrow B(x))$, or

$$(\forall x \in S)(A(x) \Longrightarrow B(x))\,. \tag{7.43}$$

- Then, you bring into your proof an object a for which you know that this object satisfies Property A, that is, you know that

$$A(a)\,. \tag{7.44}$$

- Then you derive the conclusion that $B(a)$ is true, in two steps:

  Step 1: Use the specialization rule to go from (7.43) to

$$A(a) \Longrightarrow B(a)\,. \tag{7.45}$$

  Step 2: Use Modus Ponens to go from (7.45) and (7.44) to

$$B(a)\,. \tag{7.46}$$

This combination is used all the time in proofs. The reason is that many theorems in Mathematics are of the form: "whenever something is true of an object, then something else is also true of that object", that is

$$(\forall x)(A(x) \Longrightarrow B(x)) . \tag{7.47}$$

And what you often do in proofs is take one of those theorems and apply it to a particular situation. And this is exactly what the "for all...implies" one-two punch does.

Here are some examples:

1. Take the statement that "Every positive real number has a real square root", which translates into

$$(\forall x \in \mathbb{R})(x > 0 \Longrightarrow (\exists y \in \mathbb{R})y^2 = x) .$$

   This is exactly of the form (7.47), with "$x > 0$" in the role of $A(x)$, and "$(\exists y \in \mathbb{R})y^2 = x$" in the role of $B(x)$.

   Then you can prove that 2 has a square root, by applying the "for all ... implies" one-two punch, with $a = 2$, and getting "$(\exists y \in \mathbb{R})y^2 = 2$".

2. Suppose you know that "If x is a positive real number then $x + \frac{1}{x} \geq 2$", that is, in formal language,

$$(\forall x \in \mathbb{R})(x > 0 \Longrightarrow x + \frac{1}{x} \geq 2) .$$

   (We will prove this later.) Suppose you have a real number a, and have proved that a is positive (that is, $a > 0$). Then you can draw the conclusion that $a + \frac{1}{a} \geq 2$ by using the "for all...implies" one-two punch, as follows:

   1. $(\forall x \in \mathbb{R})(x > 0 \Longrightarrow x + \frac{1}{x} \geq 2)$ [Fact proven before]

   2. $a > 0$. [Known]

   3. $a > 0 \Longrightarrow a + \frac{1}{a} \geq 2$. [Rule $\forall_{use}$, from Step 1]

   4. $a + \frac{1}{a} \geq 2$. [Rule $\Longrightarrow_{use}$, from Steps 2,3]

7.5 Using a disjunction

The rule for using a disjunction, that we are going to call "Rule $\vee_{use}$", as you may have guessed, is extremely important. It is also called the "proof by cases rule", and is one of the most widely used rules in theorem proving.

Before I state the rule, let us look at an example.

Example 14. Suppose you know that a real number x is not equal to zero, and you want to conclude that $0 < x^2$. You could reason as follows.

There are two possibilities: $0 < x$ or $0 < x$. We consider each of these two possibilities separately.

> First we assume that $\boxed{0 < x.}$
>
> Then we use Axiom OA5, which tells us that we can multiply both sides of an inequality by a positive number. Since x is positive, because we are assuming that it is, we can multiply both sides of "$0 < x$" by x, and get $x.0 < x.x$.
>
> But $x \cdot 0 = 0$ by Theorem 3.
>
> And $x \cdot x = x^2$. (See Definition 17, on page 56.)
>
> So $\boxed{0 < x^2}$.
>
> Next we assume that $\boxed{x < 0.}$
>
> Then we use Axiom OA4, which tells us that we can add a real number to both sides of an inequality. So we add $-x$ to both sides of "$x < 0$" and get $0 < -x$.
>
> Then we use Axiom OA5, which tells us that we can multiply both sides of an inequality by a positive number. Since $-x$ is positive, because we have proved that it is, under the assumption that $x < 0$, we can multiply both sides of "$0 < -x$" by $-x$, and get $(-x).0 < (-x).(-x)$.
>
> But $x \cdot 0 = 0$ by Theorem 3.
>
> And $(-x) \cdot (-x) = x \cdot x$, by Theorem 13.
>
> So $0 < x \cdot x$.

And $x \cdot x = x^2$, by Definition 17 on page 56.

So $\boxed{0 < x^2}$ in this case as well.

So we have analyzed each of the two possibilities $0 < x$ and $x < 0$, and in each case we arrived a the same conclusion, namely, that $0 < x^2$.

Hence we have proved that $\boxed{\boxed{0 < x^2}}$.

The rule for using a disjunction (Rule $\vee_{use}$, a.k.a. the proof by cases rule)

If P and Q are sentences, and you have proved $P \vee Q$ in a previous step, and then you prove another sentence R both assuming P and assuming Q, then you can go to R.

8 More definitions, and theorems with proofs

In the list of basic concepts of real number theory, "less than or equal to", "greater than", and "greater than or equal to", do not appear. And the list of basic symbols does not contain "$\leq$", "$>$", and "$\geq$". So we have to *introduce* those concepts and symbols, by explaining what they mean. This is quire easy to do.

Definition 14. If x and y are real numbers, we say that x is <u>less than or equal to</u> y, and write "$x \leq y$", if $x < y$ or $x = y$. □

Definition 15. If x and y are real numbers, we say that x is <u>larger than</u> y (or <u>greater than</u> y), and write "$x > y$", if $y < x$. □

Definition 16. If x and y are real numbers, we say that x is <u>larger than or equal to</u> y (or <u>greater than or equal to</u> y), and write "$x \geq y$", if $y < x$ or $x = y$. □

Remark 8. We could also have defined "larger than or equal to" by saying: if x and y are real numbers, we say that x is <u>larger than or equal to</u> y (or <u>greater than or equal to</u> y), and write "$x \geq y$", if $y \leq x$. Why didn't we do that? It's just a matter of taste: I chose to formulate Definition 11 as I did because it only uses the basic symbols, whereas if we say that "$x \geq y$" means "$y \leq x$", then we are using "$\leq$", which is not a basic symbol. Personally, I always prefer to give a definition in terms of the basic symbols only, whenever this is possible without much extra effort. $\square$

Now that we have defined "$\leq$", "$>$", and "$\geq$", we can prove some facts about them. These facts are so simple that we are not going to dignify them by calling them "theorems". We will just call them "propositions".

For example, we will need the transitive law for $\leq$ ("if $x \leq y$ and $y \leq z$ then $x \leq z$"). And there are also two "mixed" versions: "if $x \leq y$ and $y < z$ then $x < z$", and "if $x < y$ and $y \leq z$ then $x < z$". These are all trivial results, and the proofs are very simple, but I will give them to you, with some proofs (but not all the proofs).

I will give you the first proof in full detail, so you will see what a real, complete proof, looks like. And then I will give you the short version, which is the one that normal people would write. And then I will ask you to do the other two proofs.

Proposition 1. *If x, y, z are real numbers, then*

(1) If $x \leq y$ and $y < z$ then $x < z$.

(2) If $x < y$ and $y \leq z$ then $x < z$.

(3) If $x \leq y$ and $y \leq z$ then $x \leq z$.

(In formal language, the above statements say:

$$
\begin{array}{ll}
(1): & (\forall x \in \mathbb{R})(\forall y \in \mathbb{R})(\forall z \in \mathbb{R})((x \leq y \wedge y < z) \Longrightarrow x < z)\,, \\
(2): & (\forall x \in \mathbb{R})(\forall y \in \mathbb{R})(\forall z \in \mathbb{R})((x < y \wedge y \leq z) \Longrightarrow x < z)\,, \\
(3): & (\forall x \in \mathbb{R})(\forall y \in \mathbb{R})(\forall z \in \mathbb{R})((x \leq y \wedge y \leq z) \Longrightarrow x \leq z)\,.\ )
\end{array}
$$

A DISGUSTINGLY DETAILED PROOF OF (1).

Let x, y, z be arbitrary real numbers.

We want to prove that

$$(x \leq y \wedge y < z) \Longrightarrow x < z\,. \tag{8.48}$$

Assume that

(8.49) $$x \le y \wedge y < z\,.$$

Then $x \le y$. [Rule $\wedge_{use}$, from (8.49)]

Also, $y < z$. [Rule $\wedge_{use}$, from (8.49)]

But $x < y \vee x = y$. [By the definition of $\le$]

COMMENT: Since we know that $x < y \vee x = y$, we can do a proof by cases: We consider first the case when $x < y$, and then the case when $x = y$, and in both cases we prove that $x < y$.

Assume that $\boxed{x < y}$.

Then $x < y \wedge y < z$. [By Rule $\wedge_{prove}$, since $x < y$ and $y < z$]

But $(x < y \wedge y < z) \Longrightarrow x < z$.

[From Axiom OA3, using Rule $\forall_{use}$]

Then, since we know that $x < y \wedge y < z$, we can conclude, using Axiom OA3, that $\boxed{x < z}$.

Now assume that $\boxed{x = y}$.

Then we can use Rule SEE to substitute x for y in "$y < z$" and conclude that $\boxed{x < z}$.

So we have proved that $x < z$ in each of the two cases, $x < y$ and $x = y$, and from it follows by Rule $\vee_{use}$ that $\boxed{x < z}$.

We have proved that $x < z$ assuming that $x \le y \wedge y < z$. Hence Rule $\Longrightarrow_{prove}$ enables us to conclude that $\boxed{(x \le y \wedge y < z) \Longrightarrow x < z}$.

Since we have proved that $(x \le y \wedge y < z) \Longrightarrow x < z$ for arbitrary real numbers x, y, z, we can invoke Rule $\forall_{prove}$ and conclude that

(8.50) $$(\forall x \in \mathbb{R})(\forall y \in \mathbb{R})(\forall z \in \mathbb{R})(x \le y \wedge y < z) \Longrightarrow x < z\,,$$

which is the desired conclusion. **Q.E.D.**

A MUCH SHORTER VERSION OF THE PROOF OF (1).

Assume that $x \leq y \wedge y < z$.

Then in particular $x \leq y$, so $x < y \vee x = y$. (The words "in particular" tell us that we are really applying Rule $\wedge_{prove}$, even if we don't say so explicitly.)

And we also have $y < z$.

We now consider the two cases: $x < y$ and $x = y$.

First assume that $\boxed{x < y}$.

Since we know that $y < z$, we have $x < y \wedge y < z$, so Axiom OA3 implies that $\boxed{x < z}$.

Now assume that $\boxed{x = y}$.

Then we can use Rule SEE to substitute x for y in the formula $y < z$, and conclude that $\boxed{x < z}$.

So we have proved that $x < z$ in both the case when $x < y$ and the case when $x = y$.

This proves that

(8.51) $$(\forall x \in \mathbb{R})(\forall y \in \mathbb{R})(\forall z \in \mathbb{R})(x \leq y \wedge y < z) \Longrightarrow x < z\,,$$

which is the desired conclusion.

Q.E.D.

PROOF OF PARTS (2) AND (3). We have to prove that

(8.52) $$(\forall x \in \mathbb{R})(\forall y \in \mathbb{R})(\forall z \in \mathbb{R})((x < y \wedge y \leq z) \Longrightarrow x < z)\,,$$

and

(8.53) $$(\forall x \in \mathbb{R})(\forall y \in \mathbb{R})(\forall z \in \mathbb{R})\Big((x \leq y \wedge y \leq z) \Longrightarrow x \leq z\Big)\,.$$

YOU DO THESE TWO PROOFS. (Short versions are O.K.)

Problem 4. *Prove Formulas (8.52) and (8.53).* □

8.1 More theorems with proofs

Can we prove that 2 is not equal to zero? We know that $1 \neq 0$, because we have an axiom that says so[16].

8.1.1 An example of a very wrong proof

Here is how several Math 300 students have proved that $2 \neq 0$ in previous years.

A BAD PROOF THAT $2 \neq 0$:

By Axiom FA11,

$$1 \neq 0 \,.$$

And, of course, we can write

$$1 \neq 0$$

again.

Now let us add the two inequalities. We get

$$1 + 1 \neq 0 + 0 \,,$$

that is

$$2 \neq 0$$

(because $1 + 1 = 2$ and $0 + 0 = 0$). **Q.E.D.**

This proof is completely wrong, and would get zero points if a student wrote it in an exam.

Problem 5. *Explain why this proof is wrong. That is, imagine that you have graded this proof and have given the author a zero, and the author then comes to talk to you and asks you what is wrong with the proof. Answer the author's question. And please do not say irrelevant things such as "I don't understand the proof". And do not bother with minutiae such as "the author did not say that, in order to go from, '$1 + 1 \neq 0 + 0$' to '$2 \neq 0$' one has to use Rule SEE to substitute 2 for $1 + 1$ and 0 for $0 + 0$."* □

[16] And we do need that axiom! Otherwise, it would not be possible to prove that $1 \neq 0$. I will tell you why later.

8.1.2 Some examples of correct proofs

Our goal in this subsection is to prove that $2 \neq 0$, and some related facts.

First of all, it turns out that **it is not possible to prove that** $2 \neq 0$ **from the field axioms only.** The reason for this is somewhat subtle[17]. Let me explain.

A system of objects (which you can call "numbers", if you wish, that satisfies the 11 field axioms is called a field[18].

There are lots of fields other than $\mathbb{R}$. And in all those fields the field axioms are true. Therefore *everything that can be proved from the field axioms is true in every field.*

But it is quite easy to construct fields where $2 = 0$. (We will do this later in the course.) It follows from this that you cannot prove from the field axioms alone that $2 \neq 0$, because if this could be proved, then it would be true in every field that $2 \neq 0$.

It follows from these observations that, in order to prove that $2 \neq 0$ we have to use heavier artillery. In fact, what we need is the order axioms.

I will now show you how to prove that $2 \neq 0$ using the order axioms. But first we need some preliminary theorems that are useful in their own right.

Theorem 9. *If x is a real number, then $(-x) + x = 0$. (In formal language: $(\forall x \in \mathbb{R})(-x) + x = 0$.)*

PROOF.

[17]You could try very hard to find a proof, and if you do that you will fail, unless you cheat, i.e., take at least one unjustifiable step, as in the example of our "bad proof that $2 \neq 0$". Could you then say that "it is not possible to prove that $2 \neq 0$ from the field axioms, because I tried very hard and didn't figure out how to do it"? **NO!** That is not a valid argument.

[18]Naturally, if $\mathbb{F}$ is a field, you have to restate the axioms as applying to $\mathbb{F}$, rather than to $\mathbb{R}$. So, for example, Axiom FA4 will read "$(\forall x \in \mathbb{F})(\forall y \in \mathbb{F})x + y = y + x$".

Let x be an arbitrary real number.
Then Axiom EA1 tells us that $(-x) + x = (-x) + x$.
And $-x = 0 - x$, by the definition of $-x$ (Definition 1).
So we may use Rule SEE to substitute $0-x$ for the second occurrence of $-x$ in "$(-x) + x = (-x) + x$", thus getting

$$(-x) + x = (0 - x) + x\,. \tag{8.54}$$

But Axiom FA7 tells us that $(0 - x) + x = 0$.
Hence, using Rule SEE, we find

$$(-x) + x = 0\,. \tag{8.55}$$

So we have proved that $(-x) + x = 0$ for an arbitrary real number x, and then Rule $\forall_{prove}$ enables us to conclude that

$$(\forall x \in \mathbb{R})(-x) + x = 0\,.$$

Q.E.D.

Theorem 10. *If x is a real number, then $-x = (-1).x$. (In formal language: $(\forall x \in \mathbb{R}) - x = (-1).x$.)*

PROOF.

Let x be an arbitrary real number.
COMMENT: The strategy we are going to use is this: we know that $(-x)+x=0$*. We will compute* $(-1).x+x$ *and will find that this is also* 0*. Then it will follow that* $(-1).x+x=(-x)+x$*, and using the cancellation law for addition (Theorem 1) we will get* $(-1).x=-x$*.*
Theorem 9 tells us that $(-x)+x=0$.
On the other hand,

$$\begin{aligned}(-1).x+x &= (-1).x+x.1\\ &= x.(-1)+x.1\\ &= x.((-1)+1)\\ &= x.0\\ &= 0\,,\end{aligned}$$

where we have used (a) the fact that $x=x.1$, (b) the fact that $(-1).x=x.(-1)$, (c) the distributive law, (d) the fact that $(-1)+1=0$, which is a special case of Theorem 7, and (e) the fact that $x.0=0$ (i.e., Theorem 3).
So $(-x)+x=0$ and $(-1).x+x=0$. Therefore $(-x)+x=(-1).x+x$. It then follows from Theorem 1 that $\boxed{(-1).x=-x}$.

Q.E.D.

Theorem 11. *If* x *is a real number, then* $-(-x)=x$*. (In formal language:* $(\forall x\in\mathbb{R})-(-x)=x$*.)*

PROOF.

Let x be an arbitrary real number.

Then, by Theorem 9, $(-x)+x=0$.

Therefore $x+(-x)=0$.

Also, Theorem 7, applied to $-x$, implies $(-(-x))+(-x)=0$.

Hence $(-(-x))+(-x)=x+(-x)$.

Using Theorem 1, we can cancel the $-x$ that occurs in both sides of the previous equation, and conclude that

$$-(-x) = x\,.$$

So we have proved that $-(-x) = x$ for an arbitrary real number x. Hence Rule $\forall_{prove}$ implies that

$$(\forall x \in \mathbb{R}) - (-x) = x\,.$$

Q.E.D.

Theorem 12. *If x and y are real numbers, then $-xy = (-x).y$. (In formal language: $(\forall x \in \mathbb{R})(\forall y \in \mathbb{R}) - xy = (-x).y$.)*

PROOF.
Let x, y be arbitrary real numbers.
Then, by Theorem 8, $-x.y = (-1).(xy)$.
Also, by Axiom FA4, $(-1).(xy) = ((-1).x).y$.
And, by Theorem 10, $-x = (-1).x$.
So $(-1).(xy) = (-x).y$.
Hence $\boxed{(-x).y = -xy.}$

Q.E.D.

Theorem 13. *If x and y are real numbers, then $(-x)(-y) = xy$. (In formal language: $(\forall x \in \mathbb{R})(\forall y \in \mathbb{R})(-x).(-y) = x.y$.)*

PROOF[19]**.**
By Theorem 12, $(-x)(-y) = -(x.(-y))$.
But $x.(-y) = (-y).x$, and Theorem 12 implies $(-y).x = -yx$.
So $x.(-y) = -xy$.
Hence $(-x)(-y) = -(-xy)$.
But Theorem 11 tells us that $-(-xy) = xy$.
$(-x)(-y) = xy$. **Q.E.D.**

Corollary 1 [20] $(-1).(-1) = 1$.

[19]From now, I will omit the starting line "Let x and y be arbitrary real numbers", as long as it is clear that this should be there

[20]A corollary of a theorem is a simple result, with a very short proof, that follows easily from the theorem.

PROOF.
Apply Theorem 13, with 1 in the role of both x and y, and get $(-1).(-1) = 1.1$.
But $1.1 = 1$.
$(-1).(-1) = 1$. **Q.E.D.**

Definition 17. Let x be a real number. The square of x is the real number x^2 given by $x^2 = x.x$. □

Theorem 14. *If x is a real number, then $x^2 \geq 0$, and if $x \neq 0$ then $x^2 > 0$.*

PROOF.
First suppose that $x \neq 0$. Then there are two possibilities, namely, $0 < x$ and $x < 0$.
If $0 < x$, then Axiom OA5 enables us to multiply both sides of the inequality "$0 < x$" by the positive number x, and get $x.0 < x.x$. But $x.0 = 0$, and $x.x = x^2$. Hence $0 < x^2$, som $\boxed{x^2 > 0}$.
If $x < 0$, then Axiom OA4 enables us to add to both sides of the inequality "$x < 0$" the real number $-x$, and get $0 < -x$. Then we can multiply both sides by the positive number $-x$, getting $(-x).0 < (-x).(-x)$. But $(-x).0 = 0$, and $(-x).(-x) = x.x = x^2$. Hence $0 < x^2$, so $\boxed{x^2 > 0}$.
So we have proved that $x^2 > 0$ in both cases, when $0 < x$ and when $x < 0$. Since $0 < x \vee 0 < x$ is true when $x \neq 0$, we have proved that if $x \neq 0$ then $x^2 > 0$.

If $x = 0$ then of course $x^2 = 0$. So $x^2 \geq 0$. **Q.E.D.**

Theorem 15. $1 > 0$

PROOF.
By Theorem 14, $1.1 > 0$, because $1 \neq 0$ by Axiom FA11.
But $1.1 = 1$.
So $1 > 0$. **Q.E.D.**

Problem 6. *We have just proved that $0 < 1$, and this implies, by Axiom OA2, that $0 \neq 1$. So we have been able to prove that $0 \neq 1$. Does that mean that we can remove Axiom FA11 from the list of Axioms, since after all we have proved that $0 \neq 1$?* □

Theorem 16. $2 \neq 0$.

PROOF.
We know from Theorem 15 that $0 < 1$.
Axiom OA4 tells us that we can add any real number to both sides of an inequality, and preserve the inequality. So, if we add -1 to both sides of "$0 < 1$", we get $1 < 1 + 1$, that is, $1 < 2$.
Since $0 < 1$ and $1 < 2$. the transitive law of $<$ (i.e., Axiom OA3) implies that $\boxed{0 < 2}$. **Q.E.D.**

9 Some inequalities

Now that we know the basic properties of $<$ (and its close relatives $\leq$, $>$, $\geq$), we are ready to prove some nontrivial inequalities.

9.1 The inequality $x + \frac{1}{x} \geq 2$

Theorem 17. *If x is a positive*[21] *real number, then $x + \frac{1}{x} \geq 2$. (In formal language: $(\forall x \in \mathbb{R})(x > 0 \Longrightarrow x + \frac{1}{x} \geq 2$.)*

PROOF.

Let x be an arbitrary real number.

Assume that $x > 0$.

We want to prove that

$$x + \frac{1}{x} \geq 2\,. \tag{9.56}$$

We will prove this by contradiction.

Assume that (9.56) is not true.

Then

$$x + \frac{1}{x} < 2\,. \tag{9.57}$$

[21] The meaning of the word "positive" was discussed in Lecture 1, in a subsection called "positive, negative, nonnegative, and nonpositive numbers". As explained there, "positive" means "> 0".

Since $x > 0$, Axiom OA5 enables you to multiply both sides of (9.57) by x, getting

$$x^2 + 1 < 2x\,. \tag{9.58}$$

Then we can add $-2x$ to both sides, and get

$$x^2 + 1 - 2x < 0\,. \tag{9.59}$$

But $x^2 + 1 - 2x = (x-1)^2$. (This is easy to prove it. Try to do it.) So

$$(x-1)^2 < 0\,. \tag{9.60}$$

But Theorem 14 tells us that

$$(x-1)^2 \geq 0\,. \tag{9.61}$$

So we have proved two contradictory facts, namely, (9.60) and (9.61). The contradiction arose from assuming that (9.56) was false. Hence (9.56) us true.

Hence we have proved (9.56) under the assumption that $x > 0$. So

$$x > 0 \Longrightarrow x + \frac{1}{x} \geq 2\,. \tag{9.62}$$

Finally, we have proved (9.62) for an arbitrary x. Hence

$$(\forall x \in \mathbb{R})(x > 0 \Longrightarrow x + \frac{1}{x} \geq 2)\,. \tag{9.63}$$

Q.E.D.

Theorem 18. *If a, b are real numbers, then*

$$ab \leq \frac{a^2+b^2}{2}\,.$$

(In formal language: $(\forall a \in \mathbb{R})(\forall b \in \mathbb{R})ab \leq \frac{a^2+b^2}{2}$.)

PROOF. YOU DO IT

Problem 7. *Prove Theorem 18.*

Problem 8. *Explain what is wrong with the following proof of Theorem 18.*

Take the inequality $ab \leq \frac{a^2+b^2}{2}$.
Multiplying both sides by 2*, we get* $2ab \leq a^2 + b^2$.
Subtracting $2ab$ *from both sides, we get*

$$0 \leq a^2 + b^2 - 2ab\,.$$

But $a^2 + b^2 - 2ab = (a - b)^2$. *So we have* $0 \leq (a - b)^2$ *, which is true.*
So the inequality checks out. **Q.E.D.**

9.2 Absolute value

We have already defined "absolute value". Here is the definition again.

Definition 18. Given a real number x, the <u>absolute value</u> of x is the number $|x|$ defined as follows:

$$(9.64)\quad |x| = x \quad \text{if } 0 < x\,,$$
$$(9.65)\quad |x| = -x \quad \text{if } x < 0\,,$$
$$(9.66)\quad |x| = 0 \quad |\text{if } x = 0\,.$$ □

Now that we know what the absolute value is, we can start —proving facts about it.

9.3 Elementary properties of the absolute value

Theorem 19. *If* x, y *are real numbers, then* $|xy| = |x|.|y|$. *(In formal language:* $(\forall x \in \mathbb{R})(\forall y \in \mathbb{R})|xy| = |x|.|y|$.*)*

PROOF.

Let x, y be arbitrary real numbers.

Then either $x \geq 0$ or $x < 0$.

Assume first that $x \geq 0$.

Then $|x| = x$ (because $x \geq 0$).

And either $y \geq 0$ or $y < 0$.

Assume first that $y \geq 0$.

Then $|y| = y$.
So $|x|.|y| = xy$.
Furthermore, $xy \geq 0$.
Hence $|xy| = xy$.
Therefore $\boxed{|x|.|y| = |xy|}$.

Now assume that $y < 0$.
Then $|y| = -y$.
So $|x|.|y| = -xy$.
Furthermore, $xy \leq 0$.
Hence $|xy| = -xy$.
Therefore, once again, $\boxed{|x|.|y| = |xy|}$.

We have proved that $|x|.|y| = |xy|$, in the case when $x \geq 0$, for the two possibilities $y \geq 0$ and $y < 0$. Hence we can conclude (still assuming that $x \geq 0$), that $\boxed{|x|.|y| = |xy|}$.

We now assume that $x < 0$.

YOU COMPLETE THIS PROOF, BY DEALING WITH THE CASE $x < 0$ AND, WITHIN IT, THE TWO SUBCASES $y \geq 0$ AND $y < 0$.

Problem 9. *Complete the proof of Theorem 19*

9.4 The triangle inequality

The *triangle inequality* is one of the most important properties of the absolute value, and is constantly used in analysis[22].

Here is the statement:

Theorem 20. *If x, y are real numbers, then $|x + y| \leq |x| + |y|$. (In formal language: $(\forall x \in \mathbb{R})(\forall y \in \mathbb{R})|x + y| \leq |x| + |y|$.)*

PROOF.
First we prove a lemma[23]:

---

[22] "Analysis" is truly advanced Advanced Calculus.

[23] A *lemma* is a statement that one proves as a preliminary step towards the proof of a theorem.

Lemma 1. *If x is a real number, then $x \leq |x|$ and $-x \leq |x|$.*

PROOF OF THE LEMMA. If $x \geq 0$ then $x = |x|$, so $x \leq |x|$, and $-x \leq 0$, so $-x \leq 0 \leq |x|$, and then $-x \leq |x|$.

So we have proved that $\boxed{x \leq |x| \land -x \leq |x|}$ in the case when $x \geq 0$.

If $x < 0$, then $|x| \geq 0$, so $x \leq 0 \leq |x|$, and then $x \leq |x|$. and $|x| = -x$, so $-x \leq |x|$ as well.

So we have proved that $\boxed{x \leq |x| \land -x \leq |x|}$ in the case when $x < 0$.

Since we have proved that $\boxed{x \leq |x| \land -x \leq |x|}$ in both cases $x \geq 0$ and $x < 0$, it follows that $\boxed{x \leq |x| \land -x \leq |x|}$. **Q.E.D.**

PROOF OF THE THEOREM.

Let x, y be arbitrary real numbers.

Applying the lemma to x, we get $x \leq |x|$.

And applying the lemma to y, we get $y \leq |y|$.

Therefore $x + y \leq |x| + |y|$.

Also, applying the lemma to x, we get $-x \leq |x|$.

And applying the lemma to y, we get $-y \leq |y|$.

Therefore $(-x) + (-y) \leq |x| + |y|$.

That is, $-(x + y) \leq |x| + |y|$.

Now, it follows easily from the definition of absolute value that the absolute value of a real number u is one of the numbers u, $-u$.

So the absolute value of $x + y$ is one of the numbers $x + y$, $-(x + y)$.

But we have seen that both $x + y$ and $-(x + y)$ are $\leq |x| + |y|$.

Since $|x + y|$ is one of the numbers $x + y$, $-(x + y)$, and both numbers are $\leq |x| + |y|$, it follows that $\boxed{|x + y| \leq |x| + |y|}$.

Q.E.D.

Problem 10. Consider the following problem: *Find an upper bound*[24] *for the absolute value of the function* f *given by* $f(x) = x^3 + 5\sin x - e^x$ *on the interval* $[0, 2]$. $\square$

SOLUTION. Using the triangle inequality, we have

$$\begin{aligned} |f(x)| &= |x^3 + 5\sin x - e^x| \\ &= |x^3 + 5\sin x + (-e^x)| \\ &\leq |x^3| + |5\sin x| + |e^x| \,. \end{aligned}$$

Now, for any $x \in [0, 2]$, we have $|x^3| \leq 8$, $|5\sin x| \leq 5$, and $|e^x| \leq e^2$.

Hence $|f(x)| \leq 8 + 5 + e^2$. Since $e^2 \leq 9$ (because $e < 3$), we have $|f(x)| \leq 8 + 5 + 9$, i.e., $\boxed{|f(x)| \leq 22 \text{ for all } x \in [0, 2].}$

Problem 11. *Explain what is wrong with the following solution to the previous problem.*

SOLUTION. We have

$$\begin{aligned} |f(x)| &= |x^3 + 5\sin x - e^x| \\ &\leq |x^3| + |5\sin x| - |e^x| \,. \end{aligned}$$

Now, for any $x \in [0, 2]$, we have $|x^3| \leq 8$, $|5\sin x| \leq 5$, and $|e^x| \leq e^2 \leq 9$.

Hence $|f(x)| \leq 8 + 5 - 9$. So $\boxed{|f(x)| \leq 4 \text{ for all } x \in [0, 2].}$

Problem 12. *Prove the following inequality: If* x, y *are real numbers, then* $||x| - |y|| \leq |x - y|$. *(First write it in formal language.) You are allowed to use all the facts proved before in these notes, and in particular the triangle inequality.*

9.5 Some inequalities involving two-dimensional vectors

Definition 19. A <u>two-dimensional vector</u> is an ordered pair (a, b) of real numbers. The set of all two-dimensional vectors is called <u>real two-dimensional space</u>, and the symbol $\mathbb{R}^2$ is used to denote this set. So, instead of saying that an object $\vec{v}$ is a two-dimensional vector, we can just say "$\vec{v} \in \mathbb{R}^2$". $\square$

[24] An <u>upper bound</u> for a function f in an interval I is a real number C such that $f(x) \leq C$ for every $x \in I$. In this problem, we are asked for an upper bound for $|f(x)|$, that is, for a number C such that $|f(x)| \leq C$ for all $x \in [0, 2]$.

Example 15. The pairs[25] $(1,2)$, $(5,-3.2)$ are two-dimensional vectors. That is, $(1,2)\in\mathbb{R}^2$ and $(5,-3.2)\in\mathbb{R}^2$. □

Definition 20. If $\vec{v}\in\mathbb{R}^2$, $\vec{w}\in\mathbb{R}^2$, and $\vec{v}=(a,b)$, $\vec{w}=(c,d)$, then the dot product (or "inner product") of $\vec{v}$ and $\vec{w}$, is the real number $\vec{v}\cdot\vec{w}$ given by

$$\vec{v}\cdot\vec{w}=ac+bd\,.$$

Notice that "$\cdot$" is a binary operation[26] on $\mathbb{R}^2$, except that, unlike the other operations on a set that we have seen so far, this one does not take values in the set $\mathbb{R}^2$ but, instead, takes values in $\mathbb{R}$. □

Definition 21. The square length of a vector $\vec{v}\in\mathbb{R}^2$ given by $\vec{v}=(a,b)$ is the number $\|v\|^2$ given by

$$\|v\|^2=\vec{v}\cdot\vec{v}\,.$$

In particular, if $a\in\mathbb{R}$, $b\in\mathbb{R}$, and $\vec{v}=(a,b)$, then

$$\|v\|^2=a^2+b^2\,. \tag{9.67}$$

Definition 22. Let x be a real number. A real square root of x is a real number y such that $y^2=x$. □

Theorem 21. *Let x be a nonnegative real number. Then x has a real square root. (In formal language: $(\forall x\in\mathbb{R})(x\geq 0\Longrightarrow)(\exists y\in\mathbb{R})y^2=x\,.$)*

PROOF. This theorem cannot be proved using the tools we have so far. It requires an important axiom for $\mathbb{R}$ called the "completeness axiom", that you will see in a course more advanced than this one.

In this course, we will just admit this. So from now on we are free to use it. □

Theorem 22. *Let x be a nonnegative real number. Then:*

1. *If $x=0$ then x has exactly one real square root, namely, zero.*
2. *If $x>0$ then*

---

[25]**Clarification**: "3.2" stands for the number "three point two", that is, $\frac{32}{10}$, not for the product "three times two".

[26]Binary operations are explained on page 1.

(i) x has exactly two real square roots, one positive and the other negative,

(ii) if y is the positive real square root of x, then the negative square root is $-y$. □

PROOF. First we prove (i). Assume that $x = 0$. In that case, 0 is a real square root of x, because $0^2 = 0 \times 0 = 0$.

Are there other real square roots of x? The answer is "no", for the following reason: suppose y is a real square root of x. Then $y^2 = x$, i.e. $y^2 = 0$. That means that $y \cdot y = 0$. But Theorem 4 tells us that if the product of two real numbers is zero, then one of the numbers must be zero. Since $y \cdot y = 0$, it follows that $y = 0$. So any real square root of x must be zero. In other words, 0 is the only real square root of x. This completes the proof of (i).

We now prove (ii). Assume that $x > 0$. In that case, Theorem 19 tells us that x does have a real square root. So let y_1 be a real square root[27].

Then $y_1^2 = x$. Therefore $y_1 \neq 0$ (because if $y+1$ was equal to zero then y_1^2 would be $0 \cdot 0$, i.e., 0, so we would have $x = 0$, contradicting our assumption gthat $x \neq 0$.

Let $y_2 = -y_1$. Then y_2 is negative if y_1 is positive, and y_2 is positive if $y_1 < 0$. So we have found two different real square roots y_1, y_2 of x, one positive and one negative.

Now we want to prove that these two roots we have found are the only ones. That is, we want to show that no other real real square root of x exists. In other words, we want to show that if y is any real square root of x, then $y = y_1 \vee y = y_2$. So let y be a real square root of x. Then $y^2 = x$. Since $x = y_1^2$, we have $y^2 = y_1^2$. So $y^2 - y_1^2 = 0$.

We now use the identity $a^2 - b^2 = (a+b)(a-b)$, with y in the role of a and y_1 in the role of b. We get

$$(y + y_1)(y - y_1) = 0\,.$$

Theorem 4 tells us that if the product of two real numbers is zero, then one of the numbers must be zero. Since $(y + y_1) \cdot (y - y_1) = 0$, it follows that

[27] Here we are anticipating a rule that we will discuss soon. It's the rule $\exists_{use}$, which says that "if there exists an object of a certain kind, then we can pick one and give it a name". For example, if we know that cows exist, then we can pick a cow and call her Clarabelle, or Clarabelle Cow, or Suzy. or Bovina. Similarly, if we know that a real square root of x exists, we can pick one and call it y_1.

either $y+y_1=0$ or $y-y_1=0$. If $y+y_1=0$ then $y=-y_1=y_2$. If $y-y_1=0$ then $y=y_1$. So we have shown that any real square root y of x must be equal to y_1 or to y_2. So there are no real square roots of x other than y_1 and y_2. This completes the proof of (ii). **Q.E.D.**

Now that we know that a nonnegative real number has one nonnegative real square root and one nonpositive real square root, we can give a special name to the nonnegative root:

Definition 23. If $x \in \mathbb{R}$ and $x \geq 0$, then the square root of x is the real square root of x which is nonnegative. We use $\sqrt{x}$ to denote the square root of x. □

Definition 24. The length of a vector[28] $\vec{v}$ is thee square root of the square length of $\vec{v}$. We write $\|\vec{v}\|$ to denote the length of $\vec{v}$, so

$$\|\vec{v}\| = \sqrt{\vec{v}\cdot\vec{v}}\,.$$

In particular, if $a \in \mathbb{R}$, $b \in \mathbb{R}$, and $\vec{v}=(a,b)$, then

$$\|v\| = \sqrt{a^2+b^2}\,. \tag{9.68}$$

Theorem 23. **(The Cauchy-Schwarz inequality[29])** *If* $\vec{v} \in \mathbb{R}^2$ *and* $\vec{w} \in \mathbb{R}^2$, *then*

$$\vec{v}.\vec{w} \leq \|v\|\cdot\|w\|\,. \tag{9.69}$$

(In formal language: $(\forall \vec{v} \in \mathbb{R}^2)(\forall \vec{w} \in \mathbb{R}^2)\vec{v}.\vec{w} \leq \|v\|\cdot\|w\|$ *.)*

Example 16. Suppose $\vec{v}=(2,3)$ and $\vec{w}=(5,4)$. Then

$$\vec{v}\cdot\vec{w} = 2\times 5 + 3\times 4 = 10+12 = 22\,.$$

And

$$\|v\|^2 = \vec{v}\cdot\vec{v} = 2\times 2 + 3\times 3 = 4+9 = 13\,,$$

while

$$\|w\|^2 = \vec{w}\cdot\vec{w} = 5\times 5 + 4\times 4 = 25+16 = 41\,.$$

[28] This definition is possible because the square length $\|v\|^2$ is a nonnegative real number, thanks to (9.67), and then Theorem 21 and Definition 17 tells us that the square root of $\|v\|^2$ exists.

[29] In Russia this is called "the Cauchy-Schwarz-Bunyakovsky inequality", or just "the Bunyakovsky inequality".

so $\|\vec{v}\| = \sqrt{13}$ and $\|\vec{w}\| = \sqrt{41}$, and then

$$\|\vec{v}\| \cdot \|\vec{w}\| = \sqrt{13} \cdot \sqrt{41}\,.$$

Inequality (9.69) says that $22 \le \sqrt{13}\sqrt{41}$, and it is easy to verify that this last inequality is true. (For example, $\sqrt{13}\sqrt{41} = \sqrt{13 \times 41} = \sqrt{533}$. And $22^2 = 484$. Since $484 < 533$, it follows that $\sqrt{484} < \sqrt{533}$, that is. $22 < \sqrt{13}\sqrt{41}$.) So (9.69) holds for this particular example. □

PROOF OF THEOREM 23: YOU DO IT.

Problem 13. *Prove Theorem 23.*

HINT FOR PROBLEM 13. Let a, b, c, d be such that $\vec{v} = (a, b)$ and $\vec{w} = (c, d)$. Rewrite (9.69) in terms of a, b, c, d. then do a proof by contradiction by assuming your inequality to be false, squaring both sides, expanding out the result, and using Theorem 18 at a key point in order to get a contradiction.

Theorem 24. *(The triangle inequality in two dimensions.) If $\vec{v} \in \mathbb{R}^2$ and $\vec{w} \in \mathbb{R}^2$, then*

$$\|\vec{v} + \vec{w}\| \le \|v\| + \|w\|\,. \tag{9.70}$$

(In formal language: $(\forall \vec{v} \in \mathbb{R}^2)(\forall \vec{w} \in \mathbb{R}^2)\|\vec{v} + \vec{w}\| \le \|v\| + \|w\|\,.$)

PROOF. Let us start by computing $\|\vec{v} + \vec{w}\|^2$. We have

$$\begin{aligned}
\|\vec{v} + \vec{w}\|^2 &= (\vec{v} + \vec{w}) \cdot (\vec{v} + \vec{w}) \\
&= (\vec{v} + \vec{w}) \cdot \vec{v} + (\vec{v} + \vec{w}) \cdot \vec{w} \\
&= \vec{v} \cdot \vec{v} + \vec{w} \cdot \vec{v} + \vec{v} \cdot \vec{w} + \vec{w} \cdot \vec{w} \\
&= \|\vec{v}\|^2 + \vec{w} \cdot \vec{v} + \vec{v} \cdot \vec{w} + \|\vec{w}\|^2 \\
&= \|\vec{v}\|^2 + \vec{v} \cdot \vec{w} + \vec{v} \cdot \vec{w} + \|\vec{w}\|^2 \\
&= \|\vec{v}\|^2 + 2\vec{v} \cdot \vec{w} + \|\vec{w}\|^2\,.
\end{aligned}$$

So $\|\vec{v} + \vec{w}\|^2 = \|\vec{v}\|^2 + 2\vec{v} \cdot \vec{w} + \|\vec{w}\|^2$.

On the other hand, the Cauchy-Schwarz inequality (i.e., Theorem 23) tells us that

$$\vec{v} \cdot \vec{w} \le \|\vec{v}\| \cdot \|\vec{w}\|\,. \tag{9.71}$$

If we multiply this by 2, we get

$$2\vec{v} \cdot \vec{w} \le 2\|\vec{v}\| \cdot \|\vec{w}\|\,, \tag{9.72}$$

and if we add $\|\vec{v}\|^2 + \|\vec{w}\|^2$ to both sides of (9.72), we find

$$\|\vec{v}\|^2 + \|\vec{w}\|^2 + 2\vec{v}\cdot\vec{w} \le \|\vec{v}\|^2 + \|\vec{w}\|^2 + 2\|\vec{v}\|\cdot\|\vec{w}\|\,. \tag{9.73}$$

The right-hand side of (9.73) is equal to $(\|\vec{v}\| + \|\vec{w}\|)^2$, and the left-hand side is equal to $\|\vec{v}+\vec{w}\|^2$. So we have shown that

$$\|\vec{v}+\vec{w}\|^2 \le (\|\vec{v}\| + \|\vec{w}\|)^2\,. \tag{9.74}$$

Taking the square root of both sides, we get

$$\|\vec{v}+\vec{w}\| \le \|\vec{v}\| + \|\vec{w}\|\,, \tag{9.75}$$

which is exactly what we were trying to prove. **Q.E.D.**

10 Analysis of proofs

"The time has come," the Walrus said,
"To talk of many things:
Of shoes—and ships—and sealing wax—
Of cabbages–and kings
And why the sea is boiling hot—
And whether pigs have wings."

L. Carroll, *Through the Looking Glass*

And now the time has come to talk about *proofs*: what a proof is, how it should be written, what the rules are governing proofs, and what you should *not* do in a proof.

We have already seen several proofs. So we are now in a position to analyze these proofs, and draw from our analysis some general conclusions about proofs.

This is not going to be a complete analysis of the notion of proof, only an introduction. I will have a lot more to say later, but here are some facts that you should become aware of right now.

1. A proof consists of **clearly identifiable steps**. It should be possible to number the steps and answer questions such as "what does Step 13 say?" (In the proofs I have given you here, I did not number the steps, but sometimes one numbers them.)

2. Each step either:

   a. **makes a precise assertion** (for example, in the proof of Theorem 6, Step 1 makes the assertion that $2 \times 2 = 2 \times 2$),

   or

   b. **introduces (that is, brings in, usually by giving it a name) a new object** into the proof (for example, in the proof of Euclid's Theorem, Step 1 introduces the set of all prime numbers, by giving it the name "S"; Step 4 introduces the list L, the number n, and the numbers $p_1, \ldots, p_n$; Step 7 introduces the number N; Step 9 introduces the number q, and so on),

   or

   c. **introduces a new assumption** (for example, in the proof of Euclid's Theorem, Step 3 introduces the assumption that the set is not infinite),

   or

   d. **makes a comment** (for example, in the proof of Euclid's Theorem, Steps 2 and 10 tell the reader that we are going to prove something).

3. In each step, certain objects are mentioned. **All those objects must have been introduced before**, so that when we get to the step we know what the symbols for these objects stand for. (For example: in the proof of Euclid's Theorem, Step 10 talks about the number q and the list L. And, indeed, if you look at the previous steps, you will see that q was introduced in Step 9, and L was introduced in Step 4, which means that, by the time the readers get to Step 10, they know what Step 10 says, because they know the meaning of all the symbols used there.)

4. In each step, certain assumptions apply. **All these assumptions must have been introduced before**, so that in each step it should be clear to the reader what the assumptions are. (For example: in the proof of Euclid's Theorem, Step 4 introduces a list of all the members of S. How is that possible? It's because in Step 10 we are working

under the assumption that S is a finite set, so it is possible to pick such a list. And how do we know that the set S is finite? We know it because in Step 3 we made the assumption that the set S is finite.)

5. **Every step must be justified or at least justifiable.** This means that the step must be accompanied by a justification (that is, an explanation of why the step is valid) or, if the author chooses to omit the justification (say, so as not to write too much), the justification must exist and be easy to figure out by whoever reads your proof.

6. The justification of a step has to be that the step is either

   i. an axiom, a definition, or a theorem proved before,

   or

   ii. a statement that follows from previous steps[30]. (**IMPORTANT REMARK.** All the axioms, definitions, and theorems proved before, count as "previous steps". The idea is that you could always begin your proof, if you wanted to, by writing all the axioms, definitions, and theorems proved before, so they would be valid, true steps. And you don't actually write them, because it would be a lot of work, but you pretend they are there.)

   or

   iii. a comment, or the introduction of a new object, or an assumption, because these do not need a justification[31]

---

[30] "Follows" is a problematic word. What does it mean to say that a statement "follows" from other statements? This is going to be explained fairly soon in great detail. (Actually, the study of what "follows" from what is what Logic is about, and a lot of this course is about Logic.) But, to give you an idea, here are a couple of examples: The statement that $2 + 2 = 2 + (1 + 1)$ follows from the statement that $2 = 1 + 1$ and the statement that $2 + 2 = 2 + 2$ by applying Rule SEE. In the proof of Theorem 6, the statements "$2+2 = 2+2$" and "$2 = 1+1$" come **before** "$2+2 = 2+(1+1)$", so "$2+2 = 2+(1+1)$" follows from the previous steps. For a second example, in the proof of Euclid's Theorem, Step 8, which asserts that N has a prime factor, follows from the theorem that says that every natural number > 1 has a prime factor, together with the trivial observation that N must be > 1, because $N = M + 1$ and $M \in \mathbb{N}$.

[31] At this point, you must be very confused. Am I saying that you can introduce any assumption you want? Am I saying that you are allowed to assume things that are not

7. When I read your proof, I should be able to do the following:

   i. determine what the steps are, so that, even if the steps are not numbered, I must be able to find, say, Step 5 or Step 23, and figure out what it says;

   ii. pick a step at random, read it, and tell

      a. whether it makes an assertion, or it introduces a new object or a new assumption, or it is a comment,

   b. what all the symbols mentioned in that step mean (either because they are well-known symbols, such as the number π, or because they have been introduced in previous steps of the proof),

   c. what assumptions are valid for that step,

   d. how the step is justified, or can be justified. (So the justification should either be there, or be easy for me to figure it out.)

8. A proof moves step by step from things we know to things that we didn't know but become known as soon as we have proved them. The result that you are trying to prove (called the "conclusion") must come at the end.

9. In a proof, it is not permited to start with the conclusion, or to assert at any point something that has not been proved. You are allowed to put in comment statements, announcing, for example, that "we are going to prove that $(\forall x \in \mathbb{R})(x > 0 \Longrightarrow x + \frac{1}{x} \geq 2)$", but you are not allowed to *assert* that $(\forall x \in \mathbb{R})(x > 0 \Longrightarrow x + \frac{1}{x} \geq 2)$.

10.1 Some solved problems and some problems for you to do

Problem 14. Analyze the proof of Theorem 3 on page 29, by doing the following:

true? **Yes! I am saying exactly that!** You introduce an assumption not because you know it's true but in order to find out what would happen if that assumption was true. And often you introduce an assumption *precisely because you suspect it is not true*, so you assume that it is true and derive a contradiction, thus showing that world in which that assumption is true is an impossible world, so in the real world the assumption isn't true. This is what a **proof by contradiction** is about.

1. Rewrite the proof numbering the steps.
2. For each step,
   - i. indicate whether the step makes an assertion, introduces an object, makes an assumption, or is a comment.
   - ii. if the step makes an assertion, indicate what that assertion is,
   - iii. if the step introduces an object, indicate who that object is,
   - iv. if the step introduces an assumption, indicate what that assumption is,
   - v. if the step makes an assertion, provide a justification (which may actually be given explicitly in the proof, or may be omitted, in which case you have to figure it out),
   - vi. list all the objects mentioned in that step, and verify that each of these objects has been introduced before,
   - vii. indicate under what assumption this step is said to be valid, and verify that all these assumptions have been made before.

SOLUTION. First, here is the proof with numbered steps.

Step 1. We are going to use Rule $\forall_{prove}$ .... *This is a comment, explaining to the reader what we are going to do.*

Step 2. Let $x \in \mathbb{R}$ be arbitrary. *This step introduces* x *by declaring it to be an arbitrary real number.*

Step 3. We apply Axiom EA1 to write

$$x.0 = x.0\,. \tag{10.76}$$

This step makes the assertion that $x.0 = x.0$. *The justification is that it follows from Axiom EA1.*

Step 4. Then we use Axiom FA9 (with 0 in the role of x), to conclude that

$$0 + 0 = 0\,. \tag{10.77}$$

This step makes the assertion that $0 + 0 = 0$. *The justification is that it follows from Axiom FA9.*

Step 5. Then we use Rule SEE to substitute $0+0$ for 0 in one of the two sides of (10.76), getting

$$x.(0+0) = x.0\,. \tag{10.78}$$

This step makes the assertion that $x.(0+0) = x.0$. The justification is that it follows from Rule SEE.

Step 6. Next we use the distributive law (Axiom FA6) to conclude that

$$x.(0+0) = x.0 + x.0\,. \tag{10.79}$$

This step makes the assertion that $x.(0+0) = x.0 + x.0$. The justification is that it follows from Axiom FA6.

Step 7. Then, using Rule SEE again, we find

$$x.0 + x.0 = x.0\,. \tag{10.80}$$

This step makes the assertion that $x.0 + x.0 = x.0$. The justification is that it follows by Rule SEE.

Step 8. But Axiom FA9 implies that

$$x.0 + 0 = x.0\,. \tag{10.81}$$

This step makes the assertion that $x.0 + 0 = x.0$. The justification is that it follows from Axiom FA9.

Step 9. Hence, using Rule SEE again, we obtain

$$x.0 + x.0 = x.0 + 0\,. \tag{10.82}$$

This step makes the assertion that $x.0 + x.0 = x.0 + 0$. The justification is that it follows by Rule SEE.

Step 10. Now we use the cancellation law of addition (Theorem 1), with $x.0$ in the role of x, $x.0$ in the role of y, and 0 in the role of x, to conclude that

$$x.0 = 0\,. \tag{10.83}$$

This step makes the assertion that $x.0 = 0$. The justification is that it follows from Theorem 1.

Q.E.D.

Regarding Items vi. and vii. in the list of things we should do in our analysis, the answer is: the arbitrary number x is introduced in Step 2, and all the steps after Step 2 talk about x. No assumptions are introduced.

Problem 15. *Analyze the proof of Theorem 1 on page 27, in exactly the same way as we analyzed the proof of Theorem 3.*

Problem 16. *Analyze the proof of Theorem 17 in exactly the same way as we analyzed the proof of Theorem 3.*