MATHEMATICS 300 — FALL 2020

Introduction to Mathematical Reasoning

INSTRUCTOR'S NOTES

H. J. Sussmann

Date of this version: September 5, 2020

Contents

1 Introduction

These notes are about ***mathematical proofs***. We are going to get started by presenting some examples of proofs. Later, after we have seen several proofs, we will discuss in general, in great detail,

- What proofs are.
- How to read proofs.
- How to write and how not to write proofs.
- What proofs are for.
- Why proofs they are important.

But first, in Sections 2 and 4, I am going to show you several examples of ***proofs***.

In each of these examples, we are going to prove a ***theorem***. Theorems have ***statements***. Each statement expresses a ***proposition***, and the fact that the statement has been proved implies that the proposition is ***true***, in which case we say that the statement is true.

So maybe it is a good idea to start by clarifying the meanings of the words "theorem", "statement", "proof", and of other related words such as "proposition", "fact", and "conclusion".

1.1 Propositions, theorems and proofs

Basically, a ***proposition*** is something that can be true or false and can be the object of belief.

In other words: ***a proposition is an expression*** P ***such that it makes sense to ask the questions:***

- ***Is*** P ***true?***
- ***Is*** P ***false?***
- ***Do you believe that*** P***?***

A ***fact*** is a true proposition.
For example,

- the following are true propositions:
  - George Washington was the first president of the United States,
  - Paris is the capital of France,
  - electrons are negatively charged particles,

  - two plus two equals four,
  - if a, b are real numbers then $(a+b)^2 = a^2 + 2ab + b^2$;

- the following are false propositions:

  - John Adams was the first president of the United States,
  - Paris is the capital of Spain,
  - electrons are positively charged particles,
  - two plus two equals five,
  - if a, b are real numbers then $(a+b)^2 = a^2 + b^2$;

- the following are propositions that I don't know if they are true or false:

  - Lee Harvey Oswald was part of a conspiracy to kill President Kennedy,
  - there is intelligent extraterrestrial life,
  - every even natural number n such that $n \geq 4$ is the sum of two prime numbers[1];

- and the following are ***not*** propositions:

  - John Adams,
  - is the capital of Spain,
  - Mount Everest,
  - the book that I bought yesterday,
  - two plus two,
  - if a, b are real numbers.

A ***proof*** of a proposition P is a logical argument[2] that establishes the truth of P by moving step by step from proposition to proposition until P is reached. The proof ends with the proposition P, which is called the ***conclusion***.

For example, let us consider the proof, given on page 27, of Euclid's theorem, that the set of prime numbers is infinite: this proof consists of several ***steps***, and the very last of these steps, i.e. the conclusion, says precisely what we were trying to prove, i.e., that *the set of prime nunbers is infinite*.

Proofs can by written in a ***language***, such as English, French, Chinese, Japanese, Spanish, etc. But in addition, there is a particular language which

[1] This proposition is called "the Goldbach conjecture"; it is an unsolved problem in Mathematics.

[2] If you are worried because it is not clear to you what a "logical argument" is, do not worry. We are going to spend the whole semester discussing logical arguments and explaining what they are and how to read them and write them, so by the end of the semester you *will* know.

is prefectly suited for writing mathematical proofs: ***formal mathematical language***.

Formal mathematical language involves ***formulas***, rather than words. For example, "$2+2=4$" is an expression in formal language, i.e., a formula.

Most of our proofs will be written in a mixture of formal mathematical language and English. For example, we will write expressions such as

(#) If a and b are real numbers then $a^2 - b^2 = (a+b)(a-b)$.

But we will also explain how to write proofs in purely formal mathematical language. (And we will discuss why having a purely mathematical language is important: one of the main reasons is that ***formal mathematical language is a universal language***, that is, a language understandable by all the mathematicallu educated people in the world[3]. Another reason is that ***formal mathematical language is completely precise***: you cannot say vague things such as "the distance between A and B is small", and this is fine, because nobody knows what "small" means, so it is better if we are not allowed to say it.)

In order to write proofs in formal language, we will have to ***learn formal language***, i.e., we will have to learn to say in formal language everthing that we now say in English or in a mixture of English and formal language. For example, the sentence (#) that we wrote above will become, in formal language,

(#) $$(\forall a \in \mathbb{R})(\forall b \in \mathbb{R})\, a^2 - b^2 = (a+b)(a-b).$$

Why are proofs important? Again, this is an issue that will be taken up later, but let me sketch the answer right away:

> ***A mathematical proof of a proposition P absolutely guarantees, with complete certainty, that P is true.***

This is so for a simple reason:

[3]For example, the formula "$2+2=4$" is the same in English, French, Chinese, or any other language.

The rules of logic are designed in such a way that one can only prove, using them, propositions that are true.

Therefore, if you write a correct proof of a proposition P, that is, a proof that obeys the rules of logic, then you can be sure that P is true.

On the other hand, if you produce a purported proof of a proposition P that is not true, then we can all be sure that your proof is incorrect, in the sense that in at least one step you violeted the rules of logic.

And, in case you ask *what are those "rules of logic" that you are talking about?* The answer is: *I am about to tell you! But it is going to take me a few weeks to tell you. And, once I have told you, you will see that the rules are very simple. But you have to be patient and allow me to get you there step by step*[4].

Furthermore, ***there is no other way to know for sure that a mathematical statement is true.***

For example, consider the statement of the first theorem in this course, that the set of prime numbers is infinite. There is no way to know for sure that this is true, other than by proving it. Computing lots of prime numbers will not do, because no matter how many millions or billions or trillions of primes you may compute, you will only have computed a finite number of them, and you will never never know whether these are all the primes, or whether there are more. The proof given below shows you that, no matter how long a list of primes numbers may be, there is always at least one prime that is not on the list. And this guarantees that there are infinitely mamy primes.

1.2 Some examples of proofs

1.2.1 Expressing an integer as a difference of two squares

Let us start with a simple question: *We are given an integer k, and we want to find integers m, n such that $m^2 - n^2 = k$. That is, we want to express k as the difference of the squares of two integers.*

For some values of k, this is easy to do. For example,

[4] It's like swimming. Once you have learned to swim, it seems simple to you. But most people need to learn to swim gradually, by first practicing floating, then exhaling under water, then kicking, then maybe doing a backstroke, treading water, and so on. And, once you have learned all that, it all looks very simple.

- If I give you $k = 5$, then you can see right away that $5 = 9 - 4$, so $\boxed{5 = 3^3 - 2^2}$.
- If I give you $k = -5$, then you can see right away that $-5 = 4 - 9$, so $\boxed{-5 = 2^3 - 3^2}$.
- If I give you, say, $k = 8$, then $8 = 9 - 1$, i.e., $\boxed{8 = 3^2 - 1^2}$.
- If I give you $k = -8$, then $-8 = 1 - 9$, i.e., $\boxed{-8 = 1^2 - 3^2}$.
- If I give you $k = 28$, then it takes more work to solve the problem, but you can find a solution: $28 = 64 - 36$, and $64 = 8^2$, $36 = 6^2$, so $\boxed{28 = 8^2 - 6^2}$.
- And if $k = 29$, then you also need work, but you can find a solution: $29 = 225 - 196$, and $225 = 15^2$, $196 = 14^2$, so $\boxed{29 = 15^2 - 14^2}$.

But there are some numbers k for which you may try and try and try, and work very hard, but will not be able to find integers m, n such that $m^2 - n^2 = k$. For example, if $k = 6$, or $k = 10$, or $k = 30$, then no matter how hard you try you will not be able to find m, n.

Let us make the situation more dramatic: suppose I am offering a prize for finding integers m, n such that $m^2 - n^2 = 30$: I will give you *one million dollars* if you solve this problem. Then you will want to solve it. You will go on, and keep trying. What would it take for you to be absolutely convinced that you must stop?

Suppose someone ***proves*** to you that those numbers m, n, do not exist. If this is proved to you, and you understand the proof, and the proof convinces you, then you will be completely sure that it is impossible to find those numbers, and you will stop looking for them.

A ***proof*** of a statement S is an argument that is so convincing that after you have seen it you will be absolutely sure that S is true.

So let us prove our first theorem:

Theorem 1. *There do not exist integers m, n such that $m^2 - n^2 = 30$.*

Proof.
Suppose it is possible to pick integers m, n such that $m^2 - n^2 = 30$.
Pick integers m, n such that $m^2 - n^2 = 30$.
Then $30 = (m+n)(m-n)$, because $m^2 - n^2 = (m+n)(m-n)$.
The integer $m - n$ is either even or odd.
Suppose that $m - n$ is even.
Then $m + n$ is even as well, because $m + n = m - n + 2n$, $2n$ is even, and the sum of two even integers is even.
Since $m - n$ is even, we can pick an integer j such that $m - n = 2j$.
Since $m + n$ is even, we can pick an integer k such that $m + n = 2k$.
Then $30 = (m+n)(m-n) = (2j)(2k) = 4jk$, so 30 is divisible by 4.
Therefore $\boxed{\text{if } m - n \text{ is even then 30 is divisible by 4}}$.
Now suppose that $m - n$ is odd.
Then $m + n$ is odd as well, because $m + n = m - n + 2n$, $2n$ is even, and the sum of an odd integer and an even integer is odd.
It follows that $(m+n)(m-n)$ is odd, because $m + n$ is odd, $m - n$ is odd, and the product of two odd integers is odd.
Therefore $\boxed{\text{if } m - n \text{ is odd then 30 is odd}}$.
Since $m - n$ is either even or odd, we can conclude from the above that $\boxed{\boxed{\text{30 is either divisible by 4 or odd}}}$.
But 30 is neither divisible by 4 nor odd.
So we have proved that the statement "30 is either divisible by 4 or odd" is both true and false, and this is a contradiction.
So the assumption that it is possible to pick integers m, n such that $m^2 - n^2 = 30$ has led us to a contradcition.
Hence $\boxed{\text{it is not possible to pick integers } m, n \text{ such that } m^2 - n^2 = 30}$.

Q.E.D.

Mathematicians like to prove statements that are as general as possible. If we look at Theorem 1, we can ask an obvious question: *is Theorem 1 only about the number* 30*, or are there other numbers for which we can prove exactly the same thing?*

If you at the proof of Theorem 1, you can see immediately that thee only thing that matters about 30 is that it is neither odd nor divisible by 4. Clearly, we can exactly the same thing for other numbers that are also

neither odd nor divisible by 4, such as, for example, 6, or 10, or 50, or 202, or 4,038.

In other words, ***we can prove a much more general theorem***:

Theorem 2. *Let a be an integer such that a is neither odd not divisible by 4. Then there do not exist integers m, n such that $m^2 - n^2 = a$.*

Proof.

Let a be an arbitrary integer such that a is neither odd nor divisible by 4.

We want to prove that it is not possible to pick integers m, n such that $m^2 - n^2 = a$.

We will prove this by contradiction.

Suppose it is possible to pick integers m, n such that $m^2 - n^2 = a$.

Pick integers m, n such that $m^2 - n^2 = a$.

Then $a = (m+n)(m-n)$, because $m^2 - n^2 = (m+n)(m-n)$.

The integer $m - n$ is either even or odd.

Suppose that $m - n$ is even.

Then $m+n$ is even as well, because $m+n = m-n+2n$, $2n$ is even, and the sum of two even integers is even.

Since $m - n$ is even, we can pick an integer j such that $m - n = 2j$.

Since $m + n$ is even, we can pick an integer k such that $m + n = 2k$.

Then $a = (m+n)(m-n) = (2j)(2k) = 4jk$, so a is divisible by 4.

Therefore $\boxed{\text{if } m - n \text{ is even then } a \text{ is divisible by 4}}$.

Now suppose that $m - n$ is odd.

Then $m+n$ is odd as well, because $m+n = m-n+2n$, $2n$ is even, and the sum of an odd integer and an even integer is odd.

It follows that $(m+n)(m-n)$ is odd, because $m+n$ is odd, $m-n$ is odd, and the product of two odd integers is odd.

Therefore $\boxed{\text{if } m - n \text{ is odd then } a \text{ is odd}}$.

Since $m-n$ is either even or odd, we can conclude from the above that $\boxed{\boxed{a \text{ is either divisible by 4 or odd}}}$.

But a is neither divisible by 4 nor odd.

So we have proved that the statement "a is either divisible by 4 or odd" is both true and false, and this is a contradiction.

So the assumption that it is possible to pick integers m, n such that $m^2 - n^2 = a$ has led us to a contradcition.

Hence $\boxed{\text{it is not possible to pick integers } m, n \text{ such that } m^2 - n^2 = a}$.

And this has been proved for an arbitrary integer a such that a is neither odd nor divisible by 4.

So our proof is complete. **Q.E.D.**

Problem 1. ***Prove*** that there do not exist integers m, n such that the equation

$$m^2 - 4n^2 = 8$$

holds. □

Problem 2. ***State and prove*** a theorem that generalizes the result of Problem 1 in exactly the same way as Theorem 2 generalizes Theorem 1.□

Problem 3. For each of the following equations:

$$m^2 - 2n^2 = 1\,, \tag{1.1}$$
$$m^2 - 2n^2 = 2\,, \tag{1.2}$$
$$m^2 - 2n^2 = 3\,, \tag{1.3}$$
$$m^2 - 2n^2 = 4\,, \tag{1.4}$$
$$m^2 - 2n^2 = 5\,, \tag{1.5}$$

either ***find*** a pair (m, n) of integers for which the equation holds, or ***prove*** that such a pair does not exist. □

1.3 A preview of the division theorem for integers

The ***division theorem for integers*** is one of the most important facts of integer arithmetic. It is stated in great detail in Section 15.2, and then proved in Section 15.3, but you should be aware of it right now, without waiting until we prove it.

Here is the statement of the division theorem:

The division theorem for integers

If a, b are integers, and $b \neq 0$, then there exist unique integers q, r such that

$$a = bq + r \text{ and } 0 \leq r < |b| .$$

In formal language, the division theorem for integers says:

$$(\forall a \in \mathbb{Z})(\forall b \in \mathbb{Z})\Big(b \neq 0 \Longrightarrow (\exists! q \in \mathbb{Z})(\exists! r \in \mathbb{Z})(a = bq + r \wedge 0 \leq r < |b|)\Big) .$$

Remark 1. The statement of the division theorem for integers involves existence and uniqueness. This is discussed in great detail in Section 9, so **YOU SHOULD READ SECTION 9 RIGHT NOW.** □

The division theorem makes it possible to introduce the concepts of ***quotient*** and ***remainder*** of the division of an integer a by a nonzero integer b.

Definition 1. *If a, b are integers, and $b \neq 0$, then the unique integers q, r such that*

$$a = bq + r \qquad \text{and} \qquad 0 \leq r < |b|$$

are called, respectively, the quotient and the remainder of the division of a by b.

We use $\mathrm{QUO}(a, b)$ *and* $\mathrm{REM}(a, b)$ *to denote the quotient and the remainder of the division of a by b.* □

It follows from Definition 1 that, if $a \in \mathbb{Z}$, $b \in \mathbb{Z}$, and $b \neq 0$, then

1. $a = b \times \mathrm{QUO}(a, b) + \mathrm{REM}(a, b)$,
2. $\mathrm{QUO}(a, b) \in \mathbb{Z}$,

3. $\mathrm{REM}(a,b) \in \mathbb{Z}$ and $0 \leq \mathrm{REM}(a,b) < |b|$,

4. if q, r are integers such that $a = bq + r$ and $0 \leq r < |b|$, then $q = \mathrm{QUO}(a,b)$ and $r = \mathrm{REM}(a,b)$.

Here are some examples:

- Let $a = 25$, $b = 5$. Then $a = b \times 5 + 0$, and $0 \leq 0 < |b|$. So

$$\mathrm{QUO}(25,5) = 4 \text{ and } \mathrm{REM}(25,5) = 0\,. \tag{1.6}$$

- Let $a = 23$, $b = 5$. Then $a = b \times 4 + 3$, and $0 \leq 3 < |b|$. So

$$\mathrm{QUO}(23,5) = 4 \text{ and } \mathrm{REM}(23,5) = 3\,. \tag{1.7}$$

- Let $a = -23$, $b = 5$. Then $a = b \times (-5) + 2$, and $0 \leq 2 < |b|$. So

$$\mathrm{QUO}(-23,5) = -5 \text{ and } \mathrm{REM}(-23,5) = 2\,. \tag{1.8}$$

- Let $a = 23$, $b = -5$. Then $a = b \times (-4) + 3$, and $0 \leq 3 < |b|$. So

$$\mathrm{QUO}(23,-5) = -4 \text{ and } \mathrm{REM}(23,-5) = 3\,. \tag{1.9}$$

- Let $a = -23$, $b = -5$. Then $a = b \times 5 + 2$, and $0 \leq 2 < |b|$. So

$$\mathrm{QUO}(-23,-5) = 5 \text{ and } \mathrm{REM}(-23,-5) = 2\,. \tag{1.10}$$

1.4 Congruence of integers

Definition 2. *Let N be a natural number, and let a, b be integers. We say that a is congruent to b modulo N, and write*

$$a \equiv b \bmod N\,,$$

if $a - b$ is divisible by N. □

Remark 2. Definition 2 uses the notion of "divisibility". This is defined in Section 2.2, so **YOU SHOULD READ SECTION 2.2 RIGHT NOW.**

Here are some examples:

- 23 is congruent to 48 modulo 5, because $48 - 23 = 25$ and 25 is divisible by 5.

- 16 is congruent to 0 modulo 4, because $16 - 0 = 16$ and 16 is divisible by 4.
- -1 is congruent to 7 modulo 8, because $-1-7 = -8$ and -8 is divisible by 8.
- If a, b are integers, then a is congruent to b nodulo 1 if and only if $a = b$. This means that congruence modulo 1 is not very interesting. For that reason, when we work with congruence modulo n normally we are interestedin the case when $n > 1$.
- The following can be proved:

Theorem 3. *Let N be a natural number and let $a \in \mathbb{Z}$. Then there exists one and only one integer r such that $0 \leq r < N$ and $a \equiv r \bmod N$, and that integer is* $\mathrm{REM}(a, N)$.

Proof. **YOU DO IT.**

Problem 4. ***Prove*** Theorem 3. □

1.4.1 Useful properties of congruences

Conguences modulo a fixed integer N can be added and multiplied like ordinary equalities. Actually, the following theorem is true:

Theorem 4. *Let N be a natural number, and let a, b, c, d be integers. Then*

(1) If $a \equiv b \bmod N$ and $b \equiv c \bmod N$, then $a \equiv c \bmod N$. (This is the ***transitive law of conguence modulo*** *N.*

(2) If $a \equiv b \bmod N$ and $c \equiv d \bmod N$, then $a + c \equiv b + d \bmod N$.

(3) If $a \equiv b \bmod N$ and $c \equiv d \bmod N$, then $ac \equiv bd \bmod N$.

(4) If $a \equiv b \bmod N$, then $-a \equiv -b \bmod N$.

Proof. Proof of (1):

Let a, b, c be integers such that $a \equiv b \bmod N$ and $b \equiv c \bmod N$.

Then $a - b$ is divisible by N and $b - c$ is divisible by N.

So we may pick integers j, k such that $a - b = Nj$ and $b - c = Nk$.

Then

$$\begin{aligned} a-c &= (a-b)+(b-c) \\ &= Nj+Nk \\ &= N(j+k)\,, \end{aligned}$$

Since $j+k \in \mathbb{Z}$, it follows that $a-c$ is divisible by N, so $a \equiv c \bmod N$.

Proof of (2):

Let a, b, c, d be integers such that $a \equiv b \bmod N$ and $c \equiv d \bmod N$.

Then $a-b$ is divisible by N and $c-d$ is divisible by N.

So we may pick integers j, k such that $a-b = Nj$ and $c-d = Nk$.

Then

$$\begin{aligned} a+c-(b+d) &= (a-b)+(c-d) \\ &= Nj+Nk \\ &= N(j+k)\,, \end{aligned}$$

Since $j+k \in \mathbb{Z}$, it follows that $(a+c)-(b+d)$ is divisible by N, so $a+c \equiv b+d \mod N$.

Proof of (3): **YOU DO IT.**

Proof of (4): **YOU DO IT.**

Problem 5. ***Prove*** (3) and (4) of Theorem 4. □

Problem 6. ***Prove or disprove*** the following statement:

(*) *If* $N \in \mathbb{N}$, $a \in \mathbb{Z}$, $b \in \mathbb{Z}$, *and* $a \equiv b \bmod N$ *then* $a^2 \equiv b^2 \bmod N^2$.

NOTE: In formal language, (*) says:

$$(\forall N \in \mathbb{N})(\forall a \in \mathbb{Z})(\forall b \in \mathbb{Z})\Big(a \equiv b \bmod N \Longrightarrow a^2 \equiv b^2 \bmod N^2\Big)\,.$$

1.5 Using congruences to test the solvability of Diophatine equations

A Diophantine equation is an equation such that only integer solutions are sought or studied.

We have already studied some Diophantine equations. For example, we studied the equation[5]

$$x^2 - y^2 = 30 \tag{1.11}$$

and proved, in Theorem 1, that Equation (1.11) has no integer solutions, that is, that Equation (1.11), regarded as a Diophantine equation, has no solutions.

We also studied the Diophantine equations $x^2 - 4y^2 = 8$ (in Problem 1), and $x^2 - 2y^2 = a$, for $a = 1$, $a = 2$, $a = 3$, $a = 4$, and $a = 5$ (in Problem 3).

I will now explain how one can use congruences to gain insights into the solvability of some Diophantine equations.

Te idea is quite simple: take a Diophantine equation such as

$$x^2 - y^2 = 30\,. \tag{1.12}$$

Suppose an integer solution of (1.12) exists. Pick integers m, n such that the pair (m, n) is a solution, that is, $m^2 - n^2 = 30$. Then for every natural number N it is true that $m^2 - n^2 \equiv 30 \bmod N$. So, if we can find a natural number N such that there do not exist integers x, y for which

$$x^2 - y^2 \equiv 30 \bmod N\,,$$

then it will follow that (1.12) does not have integer solutions either.

For Equation (1.12), this is easy to do. Let us tale $N = 4$. Then, if m is an integer, we have $m \equiv 0$, or $m \equiv 1$, or $m \equiv 2$, or $m \equiv 3$, modulo 4.

If $m \equiv 0 \bmod 4$, then $m^2 \equiv 0 \bmod 4$.

If $m \equiv 1 \bmod 4$, then $m^2 \equiv 1 \bmod 4$.

If $m \equiv 2 \bmod 4$, then $m^2 \equiv 2^2 \bmod 4$, $2^2 = 4$, and $4 \equiv 0 \bmod 4$, so $m^2 \equiv 0 \bmod 4$.

If $m \equiv 3 \bmod 4$, then $m^2 \equiv 3^2 \bmod 4$, $3^2 = 9$, and $9 \equiv 1 \bmod 4$, so $m^2 \equiv 1 \bmod 4$.

So we have shown that $m^2 \equiv 0 \bmod 4$ or $m^2 \equiv 1 \bmod 4$.

Similarly, $n^2 \equiv 0 \bmod 4$ or $n^2 \equiv 1 \bmod 4$.

Hence $m^2 - n^2$ is congruent modulo 4 to $0 - 0$, or $1 - 0$, or $0 - 1$, or $1 - 1$.

[5] In Section 1.2.1. At that point we called the variables "m" and "n", but now I am calling them "x" and "y".

That is, $m^2 - n^2$ is congruent modulo 4 to 0, or 1, or -1.

But $-1 \equiv 3 \bmod 4$, so $m^2 - n^2$ is congruent modulo 4 to 0, or 1, or 3.

But $30 \equiv 2 \bmod 4$. So $m^2 - n^2$ is ***not*** congruent to 30 modulo 4, and then $m^2 - n^2$ cannot equal 30.

This proves that there do not exist integers m, n such that $m^2 - n^2 = 30$, thus giving us a different proof of Theorem 1.

Let us us the same method for a few other examples of Diophantine equations.

Question 1. ***Prove*** *that there do not exist integers* m, n *such that* $m^2 + n^2 = 3,602,963$.

Solution: We have already shown that if n is an integer then n^2 is congruent to 0 or 1 modulo 4. So $m^2 + n^2$ is congruent to 0 or 1 or 2 modulo 4. But $3,602,963 = 3 \bmod 4$. It follows that there do not exist integers m, n such that $m^2 + n^2 = 3,602,963$. □

Question 2. ***Prove*** *that there do not exist integers* m, n *such that* $m^2 + 3n^2 = 1,604$.

Solution: We look at the solvability of our equation modulo 3. If $m \in \mathbb{Z}$, then $m = 0 \bmod 3$ or $m = 1 \bmod 3$ or $m = 2 \bmod 3$. If $m = 0 \bmod 3$, then $m^2 = 0 \bmod 3$. If $m = 1 \bmod 3$, then $m^2 = 1 \bmod 3$. If $m = 2 \bmod 3$, then $m^2 = 4 \bmod 3$, but $4 = 1 \bmod 3$, so $m^2 = 1 \bmod 3$.

So we have shown that if $m \in \mathbb{Z}$ then m^2 is congruent to 0 or to 1 modulo 3.

On the other hand, it is clear that if $n \in \mathbb{Z}$ then $3n^2 \equiv 0 \bmod 3$.

It follows that $m^2 + 3n^2$ is congruent to 0 or to 1 modulo 3.

On the other hand, $1,604 \equiv 2 \bmod 3$, so it is not possible for integers m, n to satisfy $m^2 + 3n^2 = 1,604$. □

Problem 7. ***Prove*** that the Diophantine equation $61x + 23y = 1$ has a solution. □

Problem 8. ***Prove*** that the Diophantine equation $x^2 + y^2 = 3z^2$ has a unique solution. (That is, prove that there exists one and only one triple (x, y, z) of integers such that $x^2 + y^2 = 3z^2$. In formal language, the statement to be proved is: $(\exists! x \in \mathbb{Z})(\exists! y \in \mathbb{Z})(\exists! z \in \mathbb{Z}) x^2 + y^2 = 3z^2$.) □

2 An example of a proof: Euclid's proof of the infinitude of the set of prime numbers

Our first example of a proof will be Euclid's proof that there are infinitely many prime numbers. This proof is found in Euclid's *Elements* (Book IX, Proposition 20). Euclid (who was probably born in 325 BCE and died in 270 BCE) was the first mathematician to write a large treatise where mathematics is presented as a collection of definitions, postulates, propositions (i.e., theorems and constructions) and mathematical proofs of the propositions.

2.1 What Euclid's proof is about

You probably know what a "prime number" is. (If you do not know, do not worry; I will explain it to you pretty soon.) Here are the first few prime numbers:

$$2,\ 3,\ 5,\ 7,\ 11,\ 13,\ 17,\ 19\ \ldots$$

Does the list of primes stop there? Of course not. It goes on:

$$23,\ 29,\ 31,\ 37,\ 41,\ 43,\ 47,\ 53,\ 59,\ 61\ \ldots$$

And it doesn't stop there either. It goes on:

$$67,\ 71,\ 73,\ 79,\ 83,\ 89,\ 97,\ 101,\ 103\ \ldots\ .$$

Does the list go on forever? If you go on computing primes, you would find more and more of them. And mathematicians have actually done this, and found an incredibly large number of primes.

The largest known prime

As of January, 2019, the largest known prime was

$$2^{82,589,933} - 1\,.$$

(That is, 2 multiplied by itself $82,589,933$ times, minus one.) This is a huge number! It has $24,862,048$ decimal digits.

Is it possible that the list of primes stops here, that is, that there are no primes larger than $2^{82,589,933} - 1$?

Before we answer this, just ask yourself: suppose it was indeed true that the list stops with this prime number. How would you know that? If you think about it for a minute, you will see that *there is no way to know*. You could go on looking at natural numbers larger than $2^{82,589,933} - 1$, and see if among these numbers you find one that is prime. But if you don't find any it doesn't mean there aren't any. It could just be that you haven't gone far enough in your computation, and if you went farther you would find one.

In fact, no matter how many primes you may compute, you will never know whether the largest prime you have found is indeed the largest prime there is, or there is a larger one.

Can we know in some way, other than by computing lots of primes, whether the list of primes goes on forever or there is a prime number which is the largest one?

It turns out that this question can be answered by means of **reasoning**. And, amazingly, the answer is "yes, the list of primes goes on forever"! This was discovered, in the year 300 B.C., approximately, by the great Greek mathematician Euclid. Euclid's 3,000-year old proof is a truly remarkable achievement, the first result of what we would now call "number theory", one of the most important areas of Mathematics.

Euclid's theorem says the following:

Theorem. *The set of prime numbers is infinite.*

In order to prove the theorem, we need to understand the precise meaning of the terms that occur in the statement. So I will begin by explaining the meaning of "prime number" and "infinite set".

And, in order to explain what a prime number is, we will have to explain first what we mean by "divisibility", and "factors".

2.2 Divisibility of integers; factors

If you have two integers a and b, you would like to "divide a by b", and obtain a "quotient" q, i.e., an integer q that multiplied by b gives you back a. For example, we can divide 6 by 2, and get the quotient 3. And we can divide 6 by 3, and get the quotient 2.

But it is not always possible to divide a by b. For example, if $a = 4$ and $b = 3$, then an integer q such that $3q = 4$ does not exist[6].

[6] You may say that "the result of dividing 4 by 3 is the fraction $\frac{4}{3}$". That is indeed true,

Since dividing a by b is sometimes possible and sometimes not, we will introduce some new words to describe those situations when division is possible.

Definition 3. *Let a, b be integers.*

1. *We say that b <u>divides</u> a if there exists an integer k such that*
$$a = bk\,.$$
2. *We say that a is a <u>multiple</u> of b if b is a factor of a.*
3. *We say that b is a <u>factor</u> of a if b divides a.*
4. *We say that a is <u>divisible</u> by b if b divides a.*
5. *We write*
$$b|a$$
to indicate that b divides a. □

Remark 3. As the previous definition indicates,

The following are five different ways of saying exactly the same thing:

- m divides n,
- m is a factor of n,
- n is a multiple of m,
- n is divisible by m,
- $m|n$.

□

but $\frac{4}{3}$ ***is not an integer***, and so far we are working in a world in which there are integers and nothing else. If we want $\frac{4}{3}$ to exist, we have to invent new numbers—the fractions, or "rational numbers". We are going to do that pretty soon, but for the moment, since we are working with integers only, it is ***not*** possible to divide 4 by 3 and get a quotient which is an integer.

Reading statements with the "divides" symbol "|"

The symbol "|" is read as "divides", or "is a factor of".
For example, the statement "3|6" is read as "3 divides 6", or "3 is a factor of 6". And the statement "3|5" is read as "3 divides 5", or "3 is a factor of 5". (Naturally, "3|6" is true, but "3|5" is false.)
The vertical bar of "divides" has nothing to do with the bar used to write fractions. For example, " $3|6$ ***" is the statement***[a] ***"*** 3 ***divides*** 6***', which is true. And "*** $\frac{3}{6}$ ***" is a noun phrase: it is one of the names of the number also known as "*** $\frac{1}{2}$ ***", or "*** 0.5 ***".***

[a] A statement is something we can say that is true or false. A noun phrase is something we can say that stands for a thing or person. For example, "Mount Everest", "New York City", "My friend Alice", "The movie I saw on Sunday", are noun phrases. "Mount Everest is very tall", "I live in New York City", "My friend Alice studied mathematics at Rutgers", and "The movie I saw on Sunday was very boring", are statements.

Example 1. Here are some examples illustrating the use of the word "divides" and the symbol "|":

- The following statements are true:
  1. 6 divides 6,
  2. 6|6,
  3. 6 divides 12,
  4. 6|12,
  5. 1 divides 5,
  6. 1|5,
  7. 13 divides 91,
  8. 13|91,
  9. 6 divides 0,
  10. 6|0,
  11. 6 divides -6,
  12. $6|-6$,
  13. -6 divides 6,
  14. $-6|6$,
  15. 6 divides -12,

16. $-6|12$,
17. 6 divides 0,
18. $6|0$,
19. 0 divides 0,
20. $0|0$,

- and the following statements are false:

1. 6 divides 7,
2. $6|7$,
3. 0 divides 1,
4. $0|1$,
5. 12 divides 6,
6. $12|6$,
7. -5 divides 6,
8. $-5|6$,
9. $0|6$.

2.3 What is a "prime number"

Definition 4. *A prime number is a natural number p such that*

I. $p > 1$,

II. *p is not divisible by any natural numbers other than* 1 *and* p. □

And here is another way of saying the same thing, in case you do not want to talk about "divisibility".

Definition 5. *A prime number is a natural number p such that*

I. $p > 1$,

II. *There do not exist natural numbers j, k such that $j > 1$, $k > 1$, and $p = jk$.* □

2.3.1 Why isn't 1 prime?

If you look at the definition of "prime number", you will notice that, ***for a natural number p to qualify as a prime number, it has to satisfy*** $p > 1$. In other words, ***the number 1 is not prime.*** Isn't that weird? After all, the only natural number factor of 1 is 1, so the only factors of 1 are 1 and itself, and this seems to suggest that 1 *is* prime.

Well, if we had defined a number p to be prime if p has no natural number factors other than 1 and itself, then 1 *would* be prime. But we were *very* careful not to do that. Why?

The reason is, simply, that there is a very nice theorem called the "unique factorization theorem", that says that every natural number greater than 1 either is prime or can be written as a product of primes *in a unique way.* (For example: $6 = 2 \cdot 3$, $84 = 2 \cdot 2 \cdot 3 \cdot 7$, etc.)

If 1 was a prime, then the result would not be true as stated. (For example, here are two different ways to write 6 as a product of primes: $6 = 2 \cdot 3$ and $6 = 1 \cdot 2 \cdot 3$.) And mathematicians like the theorem to be true as stated, so we have decided not to call 1 a prime[7].

If you do not like this, just keep in mind that we can use words any way we like, as long as we all agree on what they are going to mean. If we decide that 1 is not prime, then 1 is not prime, and that's it. If you think that for you 1 is really prime, just ask yourself why and you will see that you do not have a proof that 1 is prime.

2.3.2 The prime factorization theorem

In our proof of Euclid's theorem, we are going to use the fact that every natural number (except 1) can be written as a product of prime numbers. This is a very important result in arithmetic[8], and we are going to prove it later.

The precise statement is as follows:

Theorem. (The prime factorization theorem.) *Every natural number n such that $n \geq 2$ is a product of primes.* □

[7] This is exactly the same kind of reason why Pluto is not a planet. Pluto is not a planet because astronomers have decided not to call Pluto a planet. Similarly, mathematicians have decided not to call 1 prime, and that's why 1 is not prime.

[8] Actually, many mathematicians call "The Fundamental Theorem of Arithmetic".

2.3.3 Clarification: What is a "product of primes"?

Like all mathematical ideas, even something as simple as "product of primes" requires a precise definition. Without a precise definition, it would not be clear, for example, whether a single prime such as 2 or 3 or 5 is a "product of primes".

Definition 6. *A natural number n is a product of primes if there exist*

1. *a natural number k,*

and

2. *a finite list*[9]
$$\mathbf{p} = (p_1, \ldots, p_k)$$
of prime numbers,

such that
$$n = p_1 \cdot p_2 \cdot \cdots \cdot p_k \,. \tag{2.13}$$
(If you are familiar with the product "$\prod$" notation, formula (2.13) says that $n = \prod_{i=1}^{k} p_i$.)

Notice that k can be equal to one. That is, ***a single prime, such as*** 2***, or*** 3***, or*** 23***, is a product of primes in the sense of our definition.*** □

Definition 7. *If n is a natural number, then a list $\mathbf{p} = (p_1, \ldots, p_k)$ of prime numbers such that (2.13) holds is called a prime factorization of n.* □

Example 2. The following natural numbers are products of primes:

- 7 (because 7 is prime); the list (7) is a prime factorization of 7,
- 24; (the list $(2, 2, 2, 3)$ is a prime factorization of 24, because $24 = 2 \times 2 \times 2 \times 3$),
- 309; (the list $(3, 103)$ is a prime factorization of 309);
- $3,895,207,331,689$. Here it would really take a lot of work to find the natural number k and the prime numbers $p_1, p_2, \ldots, p_k$ such that
$$3,895,207,331,689 = p_1 \cdot p_2 \cdot \cdots \cdot p_k \,.$$
But the prime factorization theorem guarantees to us, *without having to find the factorization of our number into a product of primes*, that $3,895,207,331,689$ is a product of primes. □

[9] Finite lists will be defined and discussed in great detail later in these notes.

2.4 Proofs by contradiction

Our proof of Euclid's theorem is going to be a ***proof by contradiction***

Proof by contradiction is probably the most important and most widely used of all proof strategies. So you should not only learn what proofs by contradiction are, but ***acquire the habit of always***[a] ***seriously considering the possibility of using the proof by contradiction strategy when you are trying to figure out how to do a proof.***

[a]Sure, I am exaggerating a little bit. There are quite a few direct proofs (that is, proofs that are not by contradiction). But the number of proofs by contradiction is huge.

Let me first explain what proofs by contradiction are, and then I will tell you why they are so important.

And the first thing I need to explain is what a ***contradiction*** is.

And, ir order to explain that, I have to discuss how to *negate* a sentence.

2.4.1 Negation

To ***negate*** (or ***deny***) a statement A is to assert that A is false. (Any such statement is called a *denial* of A)

So, for example, a denial of "7 is a prime number" is "7 is not a prime number". (But there are many other ways to write a denial of "7 is a prime number." For example, we could write "it is not true that 7 is a prime number", or "it is not the case that 7 is a prime number".)

The symbol "$\sim$" ("it's not true that")

The symbol "$\sim$", put in front of a statement, is used to assert that the statement is false.
So "$\sim$" stands for "it is not the case that", or "it is not true that".

Example 3. The following sentences are true:

- $\sim$ 6 is a prime number (that is, "6 is not a prime number"),
- $\sim$ 2 is an odd integer (that is, "2 is not an odd integer"),
- $\sim$(6 is even and 7 is even) (that is, "it's not true that 6 and 7 are both even").

The following sentences are false:

- $\sim$ 7 is a prime number (that is, "7 is not a prime number"),
- $\sim$ 3 is an odd integer (that is, "3 is not an odd integer"),
- $\sim$(6 is even or 7 is even) (that is, "it's not true that 6 is even or 7 is even"),
- $\sim$ 6 is even and 7 is even (that is, "6 is not even and 7 is even").

2.4.2 When is a negation true?

If A is a sentence, then

- $\sim A$ is true if A is false;
- $\sim A$ is false if A is true.

2.4.3 What is a contradiction?

The precise definition of "contradiction" is complicated, and requires some knowledge of logic. So let me give you a simplified definition that is easy to understand and is good enough for our purposes.

Temporary, simplified definition of "contradiction": A contradiction is a statement of the form "A and $\sim A$", that is, "A is true and A is not true". □

Example 4.

- The sentence "$2+2=7$" is ***not*** a contradiction. It is a false statement, of course, but not every false statement is a contradiction.
- The sentence "$2+2=7$ and $2+2=4$" is ***not*** a contradiction either. It is a false statement (because it is the conjunction of two sentences one of which is false), but that does not make it a contradiction.

- The sentence "$2 + 2 = 7$ and $2 + 2 \neq 7$" ***is*** a contradiction. because it is of the form "A and no A", with the sentence "$2 + 2 = 7$" in the role of A.
- The sentence "$n = 1$ and $n \neq 1$" is a contradiction.
- The sentence "John Adams was the first U.S. president" is false, but it ***not*** a contradiction.
- The sentence "John Adams was the first U.S. president and was the second U.S. president" is false, but it ***not*** a contradiction.
- The sentence "John Adams was the first U.S. president and was not the first U.S. president" ***is*** a contradiction. □

2.4.4 What is a proof by contradiction?

A ***proof by contradiction*** is a proof in which you start by assuming that the statement you want to prove is false, and you prove a contradiction. Once you have done that, you are allowed to conclude that the statement you are trying to prove is true.

To do a proof by contradiction, you would write something like this:

> We want to prove A.
>
> > Assume that A is false.
> > ⋮
> > $2 = 1$ and $2 \neq 1$.
> > And "$2 = 1$ and $2 \neq 1$" is a contradiction.
>
> So assuming that A is false has led us to a contradiction.
> Therefore A is true. **Q.E.D.**

WARNING

Having explained very precisely what a contradiction is, I have to warn you that mathematicians will often say things like " '$2+2=7$' is a contradiction". This is not quite true, but when a mathematician says that every mathematician will understand what is really intended.
What the person who said "'$2+2=7$' is a contradiction" really meant is something like this:

> Now that I have proved that $2+2=7$, I can easily get a contradiction from that, because we all know how to prove that $2+2\neq 7$, and then we can deduce from these two formulas the sentence "$2+2=7$ and $2+2\neq 7$", which is truly a contradiction.
> In other words, once I get to "$2+2=7$", it is clear to me, and to every mathematician, how to get to a contradiction from there, so there is no need to go ahead and do it, so I can stop here.

This is something mathematicians do very often[a]: ***once we get to a point where it is clear how to go on and finish the proof, we just stop there.***

For a beginning student I would recommend that you actually write your proof until you get a real contradiction, because this is the only way to make it clear to the person reading (and grading) your work that you do understand what a contradiction is.

[a] And not only mathematicians! In chess, once you get to a position from which it is clear that you can take your rival's King and win, you say "checkmate" and the game stops there.

WHAT DOES "ASSUME" MEAN?

"Assume" means "imagine". In order to prove that some statement S is true, we imagine that it is not true, that is, we explore an imaginary world W in which S is not true, and we prove that in this imaginary world something impossible (such as a contradiction, "A is true and A is not true") would have to happen. And from this we draw the conclusion that a world in which S is not true is impossible, so in the real world S must be true.

2.5 What is a finite set? What is an infinite set?

We now explain what a "finite set" is.

Definition 8. *Let S be a set,*

1. *We say that S is finite if there exist a natural number n and a finite list*[10]
$$\mathbf{a} = (a_1, a_2, \ldots, a_n)$$
with n entries which is a list of all the members of S. (This means: every member of S occurs in the list*; that is,* for every member x of S there exists a natural number j such that $j \leq n$ and $x = p_j$*.)*

2. *We say that S is infinite if it is not finite.* □

2.5.1 A simple lemma

A lemma is a statement that one proves in order to use it in the proof of a theorem. In our proof of Euclid's Theorem we are going to need the following lemma:

Lemma 1. *If a, b, c are integers, and c divides both a and b, then c divides $a + b$ and $a - b$.*

Proof. Since $c|a$ and $c|b$, we may write

$$a = cj \text{ and } b = ck\,, \tag{2.14}$$

where j and k are integers.

But then

$$a + b = c(j + k) \text{ and } a - b = c(j - k)\,, \tag{2.15}$$

and $j + k$ and $j - k$ are integers. So $c|a + b$ and $c|a - b$. **Q.E.D.**

2.6 The proof of Euclid's Theorem

The proof I am going to present here is not exactly Euclid's, but is based essentially on the same idea.

First, here is Euclid's result, again:

Theorem 5. *The set of prime numbers is infinite.*

And here is the proof.

[10] If you are wondering "what is a finite list?", then I can tell you two things: (1) you are asking a good question, (2) I will give you more information about "finite lists" later, on page 29.

Let S be the set of all prime numbers.
We want to prove that S is an infinite set.
We will prove this by contradiction.

> Suppose S is not infinite.
> Then S is a finite set.
> Since S is finite, we may write a finite list
>
> $$\mathbf{p} = (p_1, p_2, \ldots, p_n)$$
>
> of all the members of S, i.e., of all the prime numbers.
>
> Let $N = p_1.p_2.\cdots.p_n$. (That is, N is the product of all the entries of the list $\mathbf{p}$.)
> Let $M = N + 1$.
> Then $M \geq 2$, so by the prime factorization theorem (in section 2.3.2) M is a product $q_1.q_2.\cdots.q_k$ of prime numbers.
> Then q_1 is a prime number[11], and $\boxed{q_1 \text{ divides } M}$ (because $M = q_1 u$, if $u = q_2.q_3.\cdots.q_k$).
> On the other hand, since $\mathbf{p}$ is a list of all the prime numbers, and q_1 is a prime number, we can conclude that q_1 is one of the entries $p_1, p_2, \ldots, p_n$ of the list $\mathbf{p}$.
> So we may write
>
> $$q_1 = p_j\,,$$
>
> where j is one of the nmbers $1, 2, \ldots, n$.
> It follows that $\boxed{q_1 \text{ divides } N}$ (because p_j divides N and $q_1 = p_j$).
> Since q_1 divides M and q_1 divides N, it follows that q_1 divides $M - N$, by Lemma 1.
> But $M - N = 1$. So $\boxed{q_1 \text{ divides } 1}$.
> On the other hand, q_1 is prime. It then follows from the definition of "prime number" (Definition 4, on page 20) that $q_1 > 1$.
> Hence $q_1 \neq 1$.
> But then $\boxed{q_1 \text{ does not divide } 1}$, because the only natural number that divides 1 is 1.
> So $\boxed{q_1 \text{ divides } 1 \text{ and } q_1 \text{ does not divide } 1}$, which is a contradiction.

Hence the assumption that S is not an infinite set has led us to a contradiction.
Therefore $\boxed{S \text{ is an infinite set}}$. **Q.E.D.**

[11] All we need here is to have a prime number that divides M. We choose q_1, but we could equally well have chosen q_2, or any of the other q_j.

2.6.1 What is "Q.E.D."?

What does "Q.E.D." mean?

"Q.E.D." stands for the Latin phrase *quod erat demonstrandum*, meaning "which is what was to be proved". It is used to indicate the end of a proof.

Appendix: Finite lists

Finite lists have ***entries***. Sets have ***members***.
We can write[12] finite lists as follows:

1. First we write a left parenthesis, i.e., the symbol "(".
2. Then we write the names of the entries of the list, in order, beginning with entry number 1, then entry number 2, and so on. The entries must be separated by commas.
3. Then, finally, write a right parenthesis, i.e., the symbol ")".

And we can write finite sets as follows:

1. First we write a left brace, i.e., the symbol "{".
2. Then we write the names of the members of the set, in some order, separated by commas.
3. Then, finally, we write a right brace, i.e., the symbol "}".

WARNING

Be careful with the distinction between ***sets***, written with braces ("{" and "}") and ***lists***, written with parentheses ("(" and ")").
For example, the sentence

$$(1, 2, 3) = (3, 1, 2)$$

is false, but the sentence

$$\{1, 2, 3\} = \{3, 1, 2\}$$

is true.

Example 5.

[12] I am saying "we can write" rather than "we write" because there are other ways to write lists and sets. We will discuss those ways later.

- Here is the list **a** of the first ten natural numbers, in increasing order:

$$\mathbf{a} = (1, 2, 3, 4, 5, 6, 7, 8, 9, 10)\,. \tag{2.16}$$

- Here is the list **b** of the first ten natural numbers, in decreasing order:

$$\mathbf{b} = (10, 0, 8, 7, 6, 5, 4, 3, 2, 1)\,. \tag{2.17}$$

  And here is a list **c** of the first ten natural numbers, in a different order:

$$\mathbf{c} = (10, 1, 5, 8, 3, 2, 4, 9, 6, 7)\,. \tag{2.18}$$

These three lists are different. For example, the second entry of **a** is 2, whereas the second entry of **b** is 9 and that of **c** is 1.

Now let S be the set whose members are the first ten natural numbers. Then we can write

$$S = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}\,, \tag{2.19}$$

or

$$S = \{10, 9, 8, 7, 6, 5, 4, 3, 2, 1\}\,, \tag{2.20}$$

or, for example,

$$S = \{1, 3, 5, 7, 9, 2, 4, 6, 8, 10\}\,, \tag{2.21}$$

or

$$S = \{4, 2, 7, 8, 10, 1, 9, 3, 5, 6\}\,, \tag{2.22}$$

or even

$$S = \{4, 4, 2, 7, 7, 7, 5, 5, 5, 8, 10, 1, 9, 4, 3, 5, 6\}\,. \tag{2.23}$$

The sets S given by equations (2.19), (2.20), (2.21), (2.22), (2.23), are all the same set, even though the formulas describing them are different. What the formulas do is tell us who the members of the set are. So, for example, according to formula (2.19), 1 is a member of S, and 23 is not. And the other formulas also say that 1 is a member of S, and 23 is not.

The key facts are these:

- Two sets S, T are the same set if they have the same members, that is, if every member of S is a member of T and every member of T is member of S.

- Two lists **a**, **b** are the same if the first entry of **a** is the same as the first entry of **b**, the second entry of **a** is the same as the second entry of **b**, and so on. That is, $\mathbf{a} = \mathbf{b}$ if the j-th entry of **a** is the same as the j-th entry of **b** for every j.

Example 6. Let S be the set whose members are all the presidents of the United States, from George Washington to Donald Trump.

Let $\mathbf{a}$ be the list of all the presidents of the United States, from George Washington to Donald Trump, in chronological order, so

$$\mathbf{a} = (a_1, a_2, \ldots, a_{45}),$$

where, for $j = 1, 2, \ldots, 45$, a_j is the j-th U.S. president.

Then $\mathbf{a}$ has 45 entries. How many members does S have?

If you think that the answer is 45, think again!

It turns out that Grover Cleveland served two nonconsecutive terms as president, from 1885 to 1889 and from 1893 to 1897, and Congress decided that Cleveland would count as both the 22nd and the 24th president of the United States. So in the list $\mathbf{a}$, the 22nd entry a_{22} and the 24th entry a_{24} are equal. So the set S has in fact 44 members, even though the list $\mathbf{a}$ has 5 entries. □

2.7 An analogy: twin primes

Let me tell you about another problem, very similar to the one we have just discussed, for which the situation is completely different.

Definition 9. *A twin prime is a prime number p such that $p + 2$ is also prime.* □

Example 7. Here are the first few twin primes:

$$3,\ 5,\ 11,\ 17,\ 29,\ 41,\ 59,\ 71,\ 101,\ 107\ .$$ □

Now we can ask the same question that we asked for primes: does the list go on forever, or does it stop at some largest pair of twin primes?

In other words,

Are there infinitely many twin primes?

This looks very similar to the question whether there are infinitely many primes. And yet, the situation in this case is completely different:

Nobody knows whether there are infinitely twin primes. Mathematicians have been trying for more than 2,000 years to solve this problem, by proving that there are infinitely many twin primes, or that that there aren't, and so far they haven't been successful.

The twin prime conjecture is the statement that there are infinitely many pairs of twin primes. It was formulated by Euclid, about 2,300 years ago, and it is still an open problem.

THE LARGEST KNOWN TWIN PRIME

According to *Wikipedia*, as of September 2018, the current largest twin prime known was $2996863034895 \times 2^{1290000} - 1$, with $388,342$ decimal digits. It was discovered in September 2016. (The fact that the number $2996863034895 \times 2^{1290000} - 1$ is a twin prime means that it is prime, and the number $2996863034895 \times 2^{1290000} + 1$ is also prime.)

2.8 A surprising fact: non-twin primes

How about primes that are *not* twin?

Definition 10. *A non-twin prime is a prime number p such that $p + 2$ is not prime.* □

Example 8. Here are the first few non-twin primes:

2, 7, 13, 19, 23, 31, 37, 43, 47, 53,
61, 67, 73, 79, 83, 89, 97, 103 . □

And now we can ask, again, the same question that we asked for primes and for twin primes: does the list go on forever, or does it stop at some largest pair of twin primes?

In other words,

Are there infinitely many non-twin primes?

This looks very similar to the question whether there are infinitely many twin primes. And yet, the situation in this case is completely different: it is very easy to prove the following:

Theorem 6. *The set of non-twin primes is infinite.*

(I am asking you to do this proof. See Problem 16 below.)

2.9 Problems

Problem 9. Using the definition of "divides" (Definition 3), explain precisely why the statements "1 divides 5", "6 divides -6", "6 divides 0", and "0 divides 0" are true, and the statements "$5|6$" and "$0|6$" are false. □

Problem 10. Indicate which of the statements in the following list are true and which ones are false, and explain why. (That is, prove that the true statements are true and the false ones are false.)

1. Every integer is divisible by 1.
2. Every integer is divisible by 2.
3. Every integer is divisible by 0.
4. Every integer divides 1.
5. Every integer divides 2.
6. Every integer divides 0.

Problem 11. Express each of the following numbers

- 37,
- 28,
- 236,
- 2247,

as a product of prime numbers. □

Problem 12. Give a precise mathematical definition of "prime number". □

Problem 13. Give a precise mathematical definition of "twin prime". □

Problem 14. Give a precise mathematical definition of "finite set" and "infinite set". □

Problem 15. Give precise mathematical definitions of each of the following concepts:

- divides,
- is divisible by,
- factor (as in "is a factor of"),
- multiple (as in "is a multiple of"). □

Problem 16. ***Prove*** Theorem 6 (on page 32). □

Problem 17. ***Prove*** that if a, b, c are integers, $a|b$ and $b|c$, then $a|c$. □

Problem 18. ***Prove*** that if a, b are integers, $a|b$ and $b|a$, then $a = b$ or $a = -b$. □

Problem 19. The proof that was given in Section 2.6 of Euclid's Theorem uses the definition of "prime number" given on page 20. In this problem, we change the definition of "prime number" and use the following definition: *A prime number is a natural number p such that p is not divisible by any natural numbers other than 1 and p.* That is, we do not require p to be > 1. So according to this new definition 1 is now prime

Rewrite the proof of Euclid's Theorem given in Section 2.6 using the new definition of "prime number". (What you have to do is basically copy the proof, but making a few changes. For example, one of the steps of the proof given in Section 2.6 says "It follows from the definition of 'prime number' that $q_1 > 1$". This step is not valid now, because 1 is prime, so q_1 could be 1. You have to make some slight changes in the proof to adapt it to this new situation.) □

Problem 20. ***Prove*** that if p is a prime number and $p \neq 2$ then p is odd.

In the following problems, you may want to use the division theorem: ***If*** a, b ***are integers and*** $b \neq 0$***, then it is possible to write*** $a = bq + r$***, where*** q, r ***are integers such that*** $0 \leq r < |b|$**.** (For example: if a is an integer then we can write $a = 3q + r$ where $r = 0$ or $r = 1$ or $r = 2$.)

Problem 21. ***Prove*** that if p is a prime number such that $p + 2$ and $p + 4$ are also prime, then $p = 3$.

Problem 22.

1. ***Find*** at least ten different prime numbers p such that $p + 4$ is also prime.
2. ***Prove*** that the only prime number p such that $p + 4$ and $p + 8$ are also prime is $p = 3$.
3. ***Prove*** that there does not exist a prime number p such that $p + 4$, $p + 8$ and $p + 12$ are also prime.

Problem 23.

1. ***Find*** at least ten different prime numbers p such that $p + 6$ is also prime.
2. ***Find*** at least ten different prime numbers p such that $p+6$ and $p+12$ are also prime.
3. ***Find*** at least four[13] different prime numbers p such that $p+6$, $p+12$ and $p + 18$ are also prime.
4. ***Prove*** that there exists a unique prime number p such that $p + 6$, $p + 12$, $p + 18$ and $p + 24$ are also prime.
5. ***Prove*** that there does not exist a prime number p such that $p + 6$, $p + 12$, $p + 18$, $p + 24$ and $p + 30$ are also prime.

Problem 24.

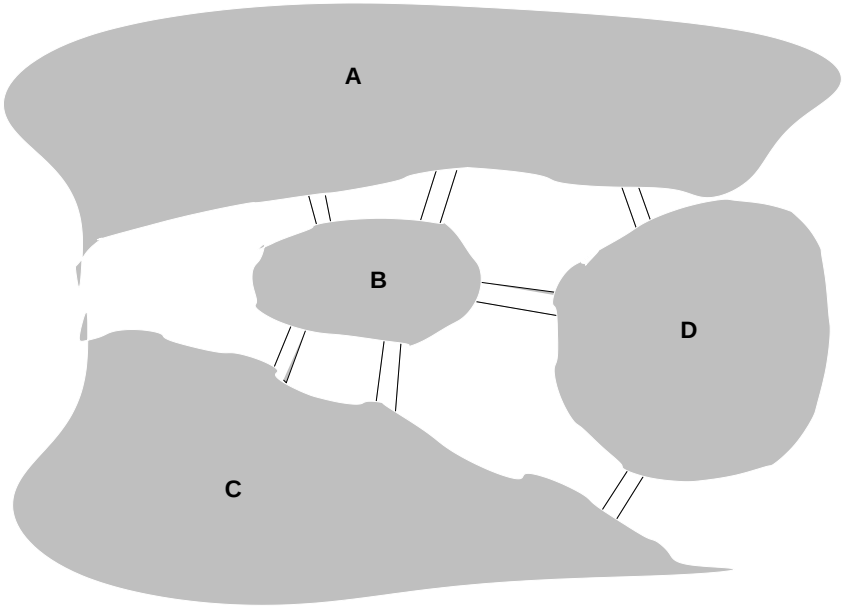
1. ***Express*** the integer 28 as a difference of two squares of integers. (That is, ***find*** two integers m, n such that $m^2 - n^2 = 28$.)
2. ***Express*** the integer 29 as a difference of two squares of integers. (That is, ***find*** two integers m, n such that $m^2 - n^2 = 29$.)
3. ***Prove*** that it is not possible to express the integer 30 as a difference of two squares of integers. (That is, ***prove*** that there do not exist two integers m, n such that $m^2 - n^2 = 30$.) □

[13]There are many more. I am just asking you to find four because I don't want to make you work too hard.

3 The seven bridges of Königsberg: a totally different example of a proof by contradiction

In 1736, the great mathematician Leonhard Euler (1707-1783) wrote a paper on the ***Königsberg bridge problem***: *Is it possible to walk through the city of Königsberg crossing each of the town's seven bridges once and only once?* The city was divided into two parts by a river crossing it, and in addition there were two islands, so the city truly had four parts, joined by seven bridges, as shown in the picture.

Euler's solution of the bridges of Königsberg problem marhed the birth of a new field of mathematics, now known as ***graph theory***, which has evolved into a major area of research, with an enormous variety of applications.



The seven bridges of Königsberg

Euler's answer was that *it is impossible to walk as proposed in the problem*: there is no way to walk through all seven bridges, crossing each bridge once and only once. Furthermore, Euler's proof is by contradiction, so it is most appropriate to include it here, to show you an example of how a proof by contradiction works.

Theorem 7. *There is no way to walk through all the seven bridges of Königsberg crossing each of the bridges once and only once.*

Proof.
We give a proof by contradiction.

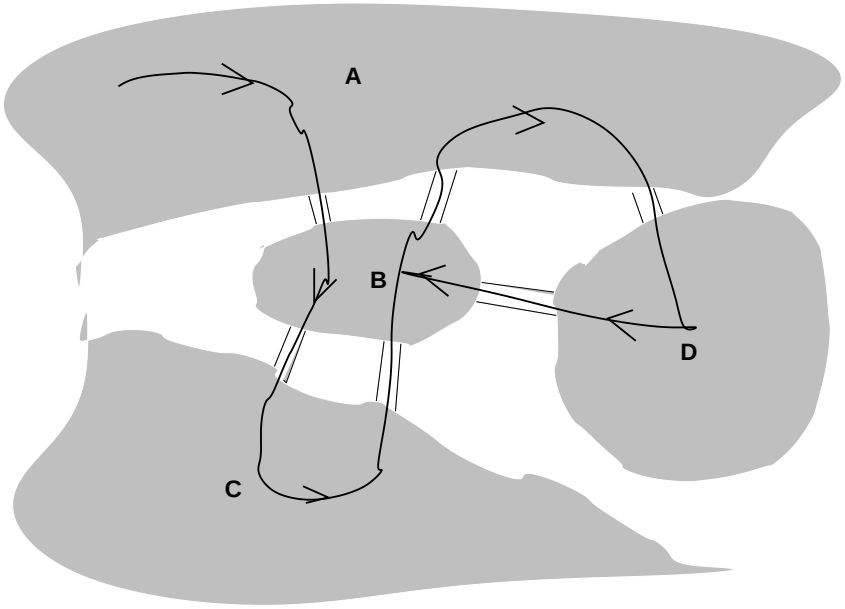
> Assume there is a way to walk through all the seven bridges, crossing each bridge once and only once.
>
> This walk starts in one of the four parts A, B, C, D into which the city is divided by the river.
>
> Call this starting part S, so S is either A or B or C or D.
>
> Furthermore, the walk ends in one of the four parts A, B, C, D.
>
> Call this part E, so E is either A or B or C or D, and E could be the same as S, or not.
>
> Let P be one of the four parts which is not S or E. (Such a part must exist, because there are four parts, and at most two of them can be S or E.)
>
> Then our walk does not start or end at P, so it must enter P at some point through one of the bridges connecting P to the other parts, and then it must leave P through a different bridge. And, if it ever enters P again, it must be through a third bridge. And then it must leave P through a fourth bridge. And so on. So the total number of bridges connecting P to other parts that are crossed by our walk has to be even, because for every bridge used to enter P there must ne a different bridge used to leave P.
>
> But our walk crosses all the bridges. And this implies that
>
> $\boxed{\text{the number of bridges connecting } P \text{ to one of the other parts is even}}$.
>
> On the other hand, for each the four parts the number of bridges connecting that part to the others is odd. (For part A the number is 3, for part B it is 5, for part C it is 3, and for part D it is also 3.)
>
> Let n be the number of bridges connecting P to one of the other parts.
>
> Then n is odd.
>
> But we have proved that n is even.
>
> So $\boxed{n \text{ is odd and } n \text{ is not odd}}$, which is clearly a contradiction.

Hence the assumption that it is possible to walk through Königsberg crossing each bridge once and only once has led us to a contradiction.

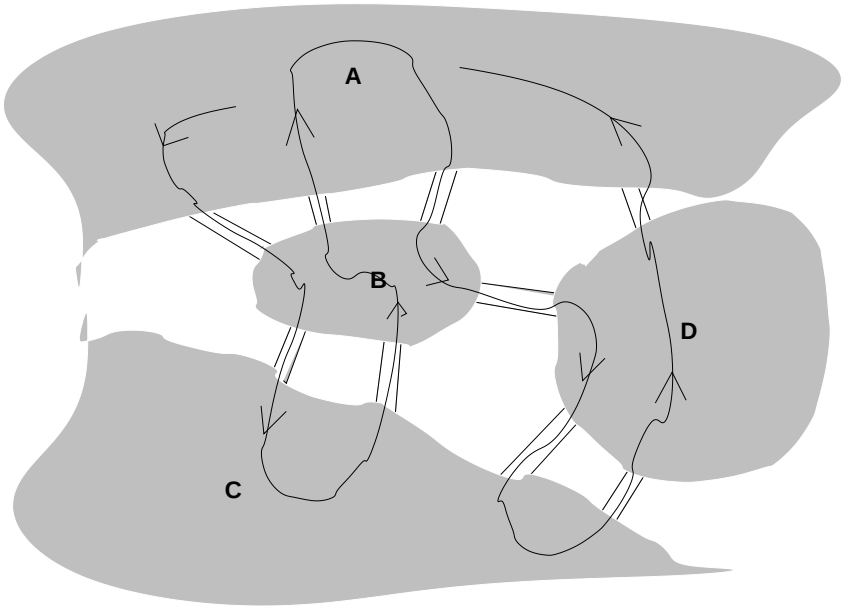
Therefore such a walk is impossible. **Q.E.D.**

3.1 Comments of the Königsberg bridge problem

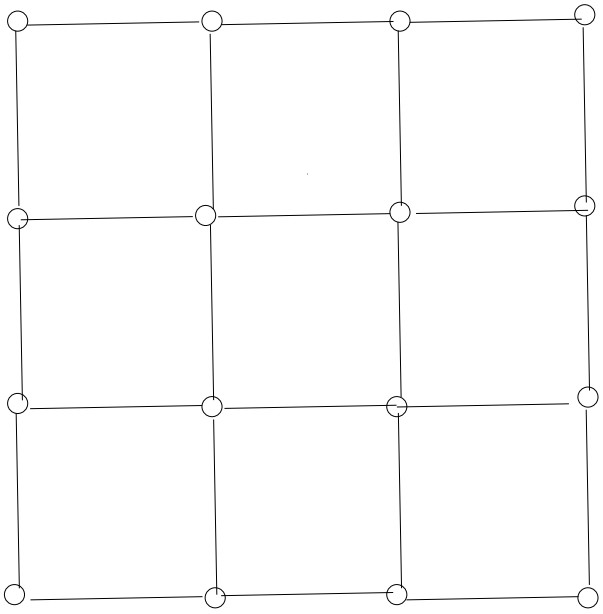
If you look at the argument given above, you see that the key point is that, if there is a way to walk through a city which is divided into several ih a number of bridges connecting



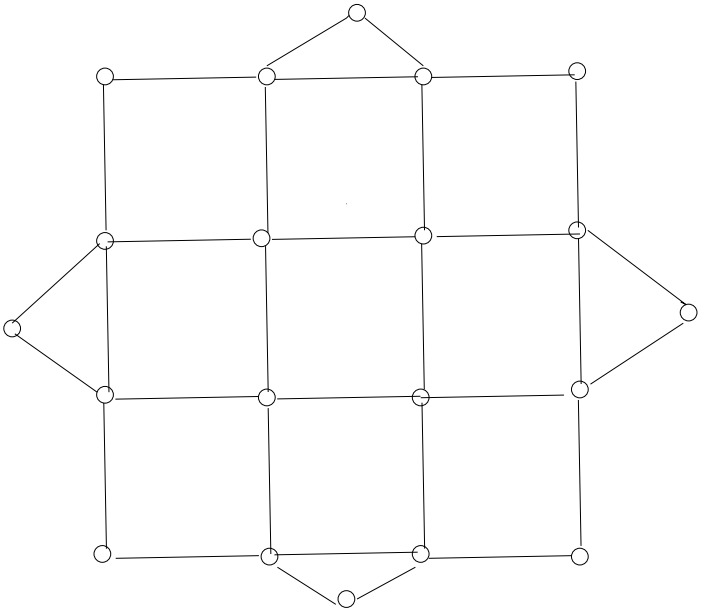
Königsberg with six bridges



Königsberg with nine bridges



16 islands and 24 bridges



20 islands and 32 bridges

4 More examples of proofs: irrationality of $\sqrt{2}$ and of other numbers

4.1 Numbers and number systems

There are several different kinds of numbers, i.e., several different number systems. It is convenient to give the number systems *names*, and to introduce mathematical symbols to represent them.

4.1.1 The most common types of numbers

Here are some examples of number systems:

- the symbol $\mathbb{N}$ stands for the set of *natural numbers*,
- the symbol $\mathbb{Z}$ stands for the set of *integers*,
- the symbol $\mathbb{Q}$ stands for the set of *rational numbers*,
- the symbol $\mathbb{R}$ stands for the set of *real numbers*,
- the symbol $\mathbb{C}$ stands for the set of *complex numbers*,
- there are sets $\mathbb{Z}_2$, $\mathbb{Z}_3$, $\mathbb{Z}_4$, $\mathbb{Z}_5$, $\mathbb{Z}_6$, and, more generally, $\mathbb{Z}_n$—the set of *integers modulo n*—for every natural number n such that $n \geq 2$. (So, for example, there are the systems $\mathbb{Z}_2$, $\mathbb{Z}_3$, $\mathbb{Z}_{10}$, $\mathbb{Z}_{11}$, $\mathbb{Z}_{5403}$.)

Some of the above kinds of numbers should be familiar to you, and others may be less so or not at all. Do not worry if you find on our list things that you have never heard of before: we will be coming back to the list later, and discussing all the items in much greater detail.

A number can belong to different number systems, in the same way as, say, a person can belong to different associations. (For example, somebody could be a member, say, of the American Association of University Professors, the Rutgers Alumni Association, and the Sierra Club. Similarly, the number 3 belongs to lots of different number systems, such as, for example, $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, and $\mathbb{R}$.)

At this point, we will just discuss $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, and $\mathbb{R}$, and we will do so very briefly. We will talk much more about these systems later, and we will also discuss later other number systems such as $\mathbb{C}$, and the $\mathbb{Z}_n$.

The symbols $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, are *special mathematical symbols*. They are *not* the capital letters N, Z, Q, R, C.

(Why do we use these special symbols? It's because mathematicians need to use lots of letters in their proofs, so they do not want to take the letters C, R, for example, and declare once and for all that they stand for "the set of all complex numbers" and "the set of all real numbers". For example, if they are working with a circle, they want to have the freedom to call the circle "C", and to say "let R be the radius of C", and this would not be allowed if the symbols "C", "R" already stood for something else. So they invented the special symbols $\mathbb{C}$, $\mathbb{R}$ to stand for the set of complex numbers and the set of real numbers, so that the ordinary letters C, R, will be available to be used as variables.)

Please do **not** say "$\mathbb{N}$ is the natural numbers", or "$\mathbb{Z}$ is the integers". When we group things together to create a set, that set is one thing, not many things. So $\mathbb{N}$ cannot be "the natural numbers". What you can, and should, say is: "$\mathbb{N}$ is the set of all natural numbers."

4.1.2 The symbol "$\in$"

If S is a set and a is an object, we write

$$a \in S$$

to indicate that a is a member of S.

And we write

$$a \notin S$$

to indicate that a is not a member of S.

How to read the "$\in$" symbol

The expression "$a \in S$" is read in any of the following ways:

- a belongs to S,
- a is a member of S,
- a is in S.

The expression "$a \notin S$" is read in any of the following ways:

- a does not belong to S,
- a is not a member of S,
- a is not in S.

Remark 4. Sometimes, "$a \in S$" is read as "a belonging to S", or "a in S", rather than "a belongs to S", or "a is in S." For example, if we write

Pick an $a \in S$,

then it would be bad English grammar to say "pick an a belongs to S". But "pick an a belonging to S", "pick an a in S", or "pick an a that belongs to S", are fine. □

Never read "$\in$" as "is contained in", or "is included in". The words "contained" and "included" have different meanings, that will be discussed later.

4.1.3 The natural numbers

The symbol $\mathbb{N}$ stands for the set of all *natural numbers*. (Natural numbers are also called "positive integers", or—sometimes—"whole numbers", or "counting numbers".) The members of this set are the numbers $1, 2, 3 \ldots$.

More precisely:

> The **natural numbers** are the numbers obtained from the number 1 by adding 1 any number of times. So, for example, the numbers 1, $1+1$ (i.e., 2), $1+1+1$ (i.e., 3), $1+1+1+1$ (i.e., 4), are natural numbers.
>
> And so are the numbers $4,503$, $46,902,444,531,322$ and $10^{10^{10^{10}}}$.
>
> The symbol $\mathbb{N}$ stands for **the set of all natural numbers**.

4.1.4 The integers

The symbol $\mathbb{Z}$ stands for the set of all *integers*.

The members of $\mathbb{Z}$ (i.e., the integers) are the natural numbers as well as 0 and the negatives of natural numbers, i.e., the numbers -1, -2, -3, etc. So, to say that a number n is an integer, we can write "$n \in \mathbb{Z}$", which we read as "n belongs to the set of integers" or, even better, as "n is an integer".

So, for example, the following statements are true:

$$\begin{aligned} 35 &\in \mathbb{N} \\ 35 &\in \mathbb{Z} \\ \sim -35 &\in \mathbb{N} \\ -35 &\in \mathbb{Z} \\ 35 &\notin \mathbb{Z} \\ 0 &\in \mathbb{Z} \\ \sim 0 &\in \mathbb{N} \\ 0 &\notin \mathbb{N} \\ 0.37 &\notin \mathbb{Z} \\ \pi &\notin \mathbb{Z} \end{aligned} \quad .$$

4.1.5 The real numbers

The symbol $\mathbb{R}$ stands for the set of all *real numbers*.

The real numbers are those numbers that you have used in Calculus. They can be positive, negative, or zero.

The positive real numbers have an "integer part", and then a "decimal expansion" that may terminate after a finite number of steps or may continue forever. (So, for example, the number 4.23 is a real number, and so is the number π. The decimal expansion of the number 4.23 terminates after two decimal figures, but the decimal expansion of π goes on forever. Here, for example, is the decimal expansion of π with 30 decimal digits:

$$3.14159265358979323846264338327 9 .$$

Using Google you can find π with one million digits. As of 2011, 10 trillion digits of π had been computed, and nobody has found any pattern! Even simple questions, such as whether every one of the digits $0, 1, 2, 3, 4, 5, 6, 7, 8, 9$ appears infinitely many times, are unresolved.)

And the negative real numbers are the negatives of the positive real numbers. So, for example, -4.23 and $-\pi$ are negative real numbers.

4.1.6 Positive, negative, nonnegative, and nonpositive numbers

In this course, "positive" means "> 0" (i.e., "greater than zero"), and "nonnegative" means "≥ 0" ("greater than or equal to zero"). So, for example, 3 and 0.7 are positive (and nonnegative), and 0 is nonnegative but not positive.

Similarly, "negative" means "< 0", and "nonpositive" means "≤ 0". So, for example, -3 and -0.7 are negative (and nonpositive), 0 is nonpositive but not negative.

4.1.7 Subsets

Definition 11. *A set A is a* **subset** *of a set B if every member of A is a member of B.*
We write "$A \subseteq B$" to indicate that A is a subset of B.

For example,

a. If, for example, S is the set of all people in the world, and T is the set of all people who live in the United States, then T is a subset of S. So the sentence "$T \subseteq S$" is true.

b. If A is the set of all animals, and G is the set of all giraffes, then G is a subset of A, so the sentence "$G \subseteq A$" is true.

c. Let S be the set of all people who live in the United States, and let C be the set of all U.S. citizens. Is C a subset of S? The answer is

"no", because there are U.S. citizens who do not live in the U.S., so these people are members of C but not of S, so it's not true that every member of C belongs to S.

And here are some mathematical examples:

I. The following sentences are true:

$$\mathbb{N} \subseteq \mathbb{Z}\,,$$
$$\mathbb{N} \subseteq \mathbb{R}\,,$$
$$\mathbb{Z} \subseteq \mathbb{R}\,,$$

because every natural number is an integer, every natural number is a real number, and every integer is a real number.

II. And the following sentences are false:

$$\mathbb{Z} \subseteq \mathbb{N}\,,$$
$$\mathbb{R} \subseteq \mathbb{N}\,,$$
$$\mathbb{R} \subseteq \mathbb{Z}\,.$$

(For example, it is not true that $\mathbb{Z} \subseteq \mathbb{N}$, because not every integer is a natural number since, for example, $0 \in \mathbb{Z}$ but $0 \notin \mathbb{N}$.)

4.1.8 The word "number", in isolation, is too vague

As we have seen, there are different kinds of numbers. So, if you just say that something is a "number", without specifying what kind of number it is, then this is too vague. In other words,

Never say that something is a "number", unless you have made it clear in some way what kind of "number" you are talking about.

For example, suppose you are asked to define "divisible", and you write:

> A number a is divisible by a number b if we can write $a = bc$ for some number c.

This is too vague! What kind of "numbers" are we talking about? Could they be real numbers?. If this was the case, then 3 would be divisible by 5, because $3 = 5.z$, if we take $z = 3/5$. But we do not want 3 to be divisible by 5. And we want the "numbers: we are talking about to be integers.

So here is a correct definition of "divisible":

Divisibility of integers: We say that an integer a is divisible by an integer b (or that a is a multiple of b, or that b is a factor of a, or that b divides a), if we can write

$$a = bc$$

for some integer c. □

For example, the following sentences are true:

$$\begin{gathered} 3 \text{ divides } 6\,, \\ -3 \text{ divides } 6\,, \\ 6 \text{ is divisible by } 3\,, \\ 6 \text{ is a multiple of } 3\,, \\ 3 \text{ is a factor of } 6\,. \end{gathered}$$

4.2 Existential statements

In the definition of divisibility given above, we have used the words "we can write". This language makes it sound as though, in order to decide whether, say, 3 divides 6, we need to have somebody there who "can write" things. This should not be necessary: "3 divides 6" would be a true sentence even if there was nobody around to do any writing. So it is much better to use a more impersonal language:

Divisibility of integers

DEFINITION. An integer a is divisible by an integer b (or a is a multiple of b, or b is a factor of a, or b divides a), if there exists an integer c such that

$$a = bc\,.$$

The sentence "there exists an integer c such that $a = bc$" is an example of nn **existential sentence**, i.e., a sentence that asserts that an object of a certain kind exists. Later, when we learn to write mathematics in formal language (that is, using only formulas), we will see that this sentence can be written as follows:

$$(\exists c \in \mathbb{Z})a = bc\,. \tag{4.24}$$

The symbol "$\exists$" is the **existential quantifier symbol**, and the expression "$(\exists c \in \mathbb{Z})$" is an **existential quantifier**, and is read as "there exists an integer c such that".

So Sentence (4.24) is read as "there exists an integer c such that $a = bc$". And it can also be read as "$a = bc$ for some integer c", or "it is possible to pick an integer c such that $a = bc$". (I recommend the "it is possible to pick ..." reading.)

4.2.1 The rule for using existential statements (Rule $\exists_{use}$)

Suppose you know that cows exist, that is that

$$(\exists x)x \text{ is a cow}. \tag{4.25}$$

Then the rule for using existential statements says that we can introduce into our conversation a cow, and give her name, by saying something like "pick a cow and call her Suzy".

In general,

- For a sentence $(\exists x)P(x)$, a witness is an object a such that $P(a)$. (For example: for the sentence (4.25), a witness is any a such that a is a cow, that is, any cow.)
- For a sentence $(\exists x \in S)P(x)$, a witness is an object a which belongs to S and is suchthat $P(a)$. (For example, if C is the set of all cows, then a witness for the sentence $(\exists x \in C)x$ is brown is any brown cow.)

The ***rule for using existential statements*** (Rule $\exists_{use}$) says that, ***if you know that an existential statement is true, then you can "pick a witness and give it a name".***

For example: suppose you know that a natural number n is not prime and is > 1. Then you know that the following is true: $(\exists m \in \mathbb{N})(m|n$ and $m \neq 1$ and $m \neq n)$. (That is, n has a factor which is a natural number and is not equal to 1 or n.) Then Rule $\exists_{use}$ says that we can pick a witness and call it a, that is, we can pick a natural number a such that $a|n$, $a \neq 1$ and $a \neq n$.

Rule $\exists_{use}$

- From
$$(\exists x)P(x)$$
you can go to "Let w be a witness for $(\exists x)P(x)$, so $P(w)$," or "Pick a witness for $(\exists x)P(x)$ and call it w", or "Pick a w such that $P(w)$."

- From
$$(\exists x \in S)P(x)$$
you can go to "Let w be a witness for $(\exists x \in S)P(x)$, so $w \in S$ and $P(w)$," or "Pick a witness for $(\exists x \in S)P(x)$ and call it w", or "Pick a w such that $w \in S$ and $P(w)$."

For example:

i. If you know that Polonius has been killed, but you do not know who did it, then you can talk about the person who killed Polonius and give a name to that person, for example, call him (or her) "the killer".

ii. if you know that an equation (say, the equation $3x^2 + 5x = 8$) has a solution (that is, you know that the existential statement "there exists a real number x such that $3x^2 + 5x = 8$" is true) then you are allowed to pick a solution and call it, for example[14], "a".

[14]Can you call this solution x? This is a complicated issue. Think of this as follows: the letter x is really a slot where you can put in a number. A number that can be put in the slot so as to make the formula true is called a "solution". The solution and the slot are two different things. So it is not a good idea to use the same name for both. If you do things *very* carefully, it turns out that it is O.K. to call both the slot and a solution with the same name, but I strongly recommend that you do not do it. For example the equation $3x^2 + 5x = 8$ has are two solutions, namely, 1 and $-\frac{8}{3}$. Which one is "x"? You cannot call both of them "x", because they are different. So I think it is better to call one of the solutions a (or A, or u, or U, or p, or P, or α, or $\heartsuit$) and then call the other one a different name (say b, or B, or v, or V, or q, or Q, or β, or $\clubsuit$).

4.3 Pythagoras' Theorem and two of its proofs

Pythagoras' Theorem is one of the oldest and most important theorems in Mathematics. It is named after the Greek mathematician and philosopher Pythagoras, who lived approximately from 570 to 495 BCE, although there is a lot of evidence that the theorem (but probably not the proof) was known before, by the ancient Babylonians.

The statement of the theorem is as follows:

Theorem 8. (Pythagoras' Theorem) *If T is a right triangle*[15]*, c is the length of the hypothenuse*[16] *of T, and a, b are the lengths of the other two sides, then*

$$a^2 + b^2 = c^2 \,. \tag{4.26}$$

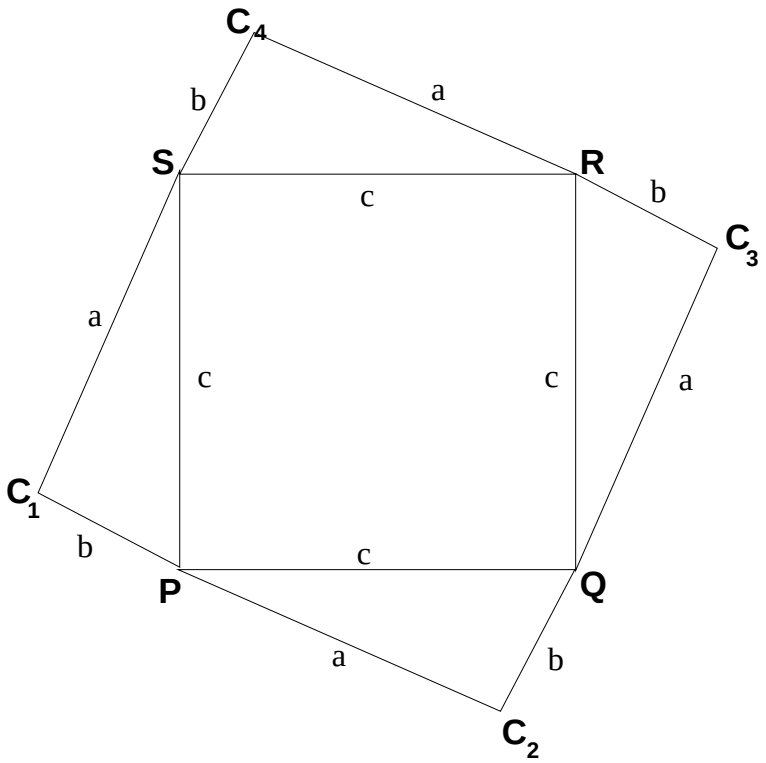
There are many different proofs of Pythagoras' Theorem. I am going to give you two proofs.

Pythagoras' proof. We draw a $c \times c$ square $PQRS$, and then attach at each side a copy[17] of T as shown in the picture.

[15] A right triangle is a triangle having one right angle

[16] The hypothenuse of a right triangle T is the side opposite to the right angle of T.

[17] For those who have studied Euclidean Geometry in high school: a copy of a figure F is a figure F' congruent to F. "Congruent to F" means: "obtainable from F by combining displacements and rotations. For example, the triangles QC_3R, RC_4S, and SC_1P are all congruent to PC_2Q.



The point P lies on the straight line segment from C_1 to C_2, because

1. If α_1 is the angle at S of the triangle SC_1P, and α_2 is the angle at P of the triangle PC_2Q, then $\alpha_1 = \alpha_2$, because the triangles SC_1P and PC_2Q re congruent.

2. Similarly, if β_1 is the angle at P of the triangle SC_1P, and β_2 is the angle at Q of the triangle PC_2Q, then $\beta_1 = \beta_2$, because the triangles SC_1P and PC_2Q are congruent.

3. Since SC_1P and PC_2Q are both right triangles, and the sum of the angles of every triangle is 180^o, we have

$$\alpha_1 + \beta_1 + 90^o = 180^o \text{ and } \alpha_2 + \beta_2 + 90^o = 180^o\,,$$

   so

$$\alpha_1 + \beta_1 = 90^o \text{ and } \alpha_2 + \beta_2 = 90^o\,.$$

4. Since $\alpha_1 = \alpha_2$, it follows that $\alpha_2 + \beta_1 = 90^o$,

5. Hence the angle θ between the segments PC_1 and PC_2 is equal to $\alpha_2 + 90^o + \beta_1$, i.e., to 180^o. This proves that the segments PC_1 and PC_2 lie on the same straight line, so P lies on the segment C_1C_2.

A similar argument shows that Q lies on the segment C_2C_3, R lies on the segment C_3C_4, and S lies on the segment C_4C_1.

So the polygonal $C_1PC_2QC_3RC_4SC_1$ is a square.

Let $d = a + b$. Then the sides of the square $C_1C_2C_3C_4$ have length d.

Therefore the area of the square $C_1C_2C_3C_4$ is d^2.

On the other hand, the smaller square $PQRS$ has side of length c, so its area is c^2. Each of the four triangles has area $\frac{ab}{2}$. So the area of $C_1C_2C_3C_4$ is equal to $c^2 + 4 \times \frac{ab}{2}$, i.e., to $c^2 + 2ab$.

It follows that

$$\begin{aligned} (a+b)^2 &= d^2 \\ &= c^2 + 4 \times \frac{ab}{2} \\ &= c^2 + 2ab\,. \end{aligned}$$

On the other hand, $(a+b)^2 = a^2 + b^2 + 2ab$. It follows that

$$a^2 + b^2 + 2ab = c^2 + 2ab\,.$$

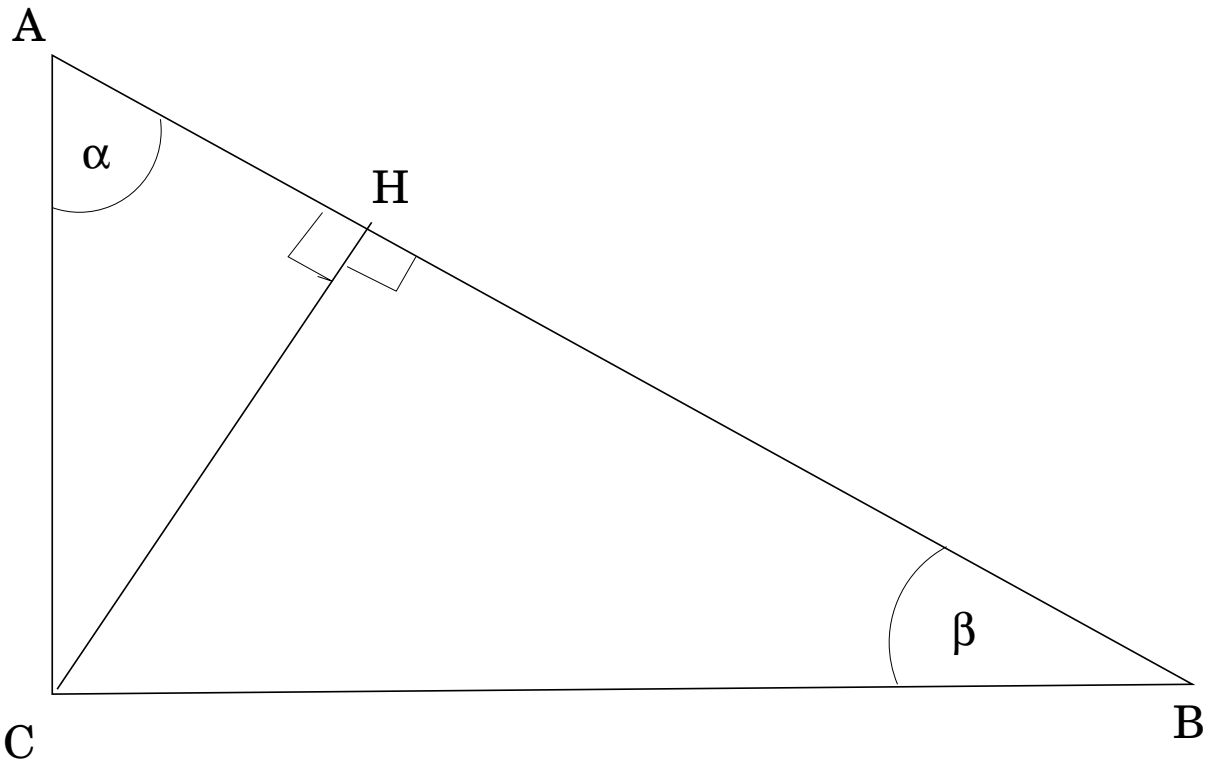
Subtracting $2ab$ from both sides, we get

$$a^2 + b^2 = c^2\,,$$

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

Proof using similar triangles. Let C be the vertex of T where the right angle is located, and let A, B be the other two vertices.

Draw a line through C perpendicular to the line AB, and let H be the point where this line intersects the line AB.



Let α, β be the angles of T at A, B, so $\alpha + \beta = 90^o$. The angle of ACH at H is also 90^o, and the angle at A is α. Hence the angle of ACH at C is β. So the triangles ABC and ACH are similar. Hence the sides opposites to equal angles are proportional. That is:

$$\frac{|AC|}{|AH|} = \frac{|AB|}{|AC|},$$

from which it follows that

$$|AC|^2 = |AH| \cdot |AB|.$$

A similar argument shows that

$$|BC|^2 = |BH| \cdot |AB|.$$

Adding both equalities we get

$$\begin{aligned} a^2 + b^2 &= |AH| \cdot |AB| + |HB| \cdot |AB| \\ &= \Big(|AH| + |HB|) \cdot |AB| \\ &= |AB| \cdot |AB| \\ &= |AB|^2 \\ &= c^2 \,. \end{aligned}$$

So $a^2 + b^2 = c^2$, as desired. **Q.E.D.**

4.4 Rational and irrational numbers

In this section we will prove a very important fact, namely, that "the number $\sqrt{2}$ is irrational". This means, roughly, the same thing as "there does not exist a rational number r such that $r^2 = 2$." (The two statements do not say exactly the same thing. I will discuss how they differ later.)

But first I want to explain what this means and why this result is so important. And to do this we need a small philosophical digression into the question: ***what is a "number"?***. (If you are not interested in philosophical questions, you may skip this discussion and move on to subsection 4.4.4.)

4.4.1 What are "numbers"?

We have already been talking quite a bot about "numbers", but I never told you what a "number" is. The question "what is a number?" is not an easy one to answer, and I will not even try. But there are some tings that can be said.

1. ***Numbers*** are, basically, tags (or labels) that we use to specify the amount or quantity of something, i.e., to answer the questions "how much ...?" or "how many ...?"

2. Since ancient times, it was understood that there are at least two kinds of "numbers":

   (a) The ***counting numbers***, that we use to specify amounts of discrete quantities, such as coins, people, animals, stones, books, etc.

   - counting numbers are used to ***count***: 1, 2, 3, 4, 5, and so on,
   - they are the ones that ***answer questions of the form "how many ... are there?"***;

- they ***vary in discrete steps***: they start with the number 1, then they "jump" from 1 to 2, and there is no other counting number between 1 and 2, then they "jump" from 2 to 3, and there is no other counting number between 2 and 3, and so on.

(b) The ***measuring numbers***, that we use to specify amounts that can vary continuously, such as lengths, areas, volumes, weights.

- measuring numbers are used to ***measure*** continuously varying quantities;
- they are the ones that ***answer questions of the form "how much ... is there?"***;
- they ***vary continuously***, so that, for example, when you pour water into a cup, if at some time point there are 10 ounces in the cup, and later there are 12 ounces, it does not occur to us that the amount of water in the cup may have jumped directly from 10 to 12 ounces: we understand that at some intermediate time there must have been 11 ounces, and at some time before that there must have been 10.5 ounces, and at some time before that there must have been 10.25 ounces, and at some time before the amount of water in the cup was 10.153098341832189504 82 ounces; and so on[18]. At no time did the amount of water "jump"[19] from some value u to some larger value v.
- they ***can be subdivided indefinitely***: for example
  - You can take a segment of length 1 (assuming we have fixed a unit of length), and divide it into seven equal segments, each one of which has length $\frac{1}{7}$. And then you can draw segments whose lengths are $\frac{3}{7}$, or $\frac{4}{7}$, or $\frac{9}{7}$, or $\frac{23}{7}$, thus getting fractional lengths.

[18] WARNING: The words "and so on" here are very imprecixse. It's not at all what they mean. When I talk about the counting numbers and I write "1, 2, 4, 5, and so on", you know exacrtly what comes next: it's 6. But when I write "11, 10.5, 10.25, 10.153098341832189504 82, and so on", I haven't the faintest idea what comes next! So the "and so on" for counting numbers is acceptable, but the "and so on" for measuring numbers is not, and when we do things rigorously and precisely we must get rid of it.

[19] To make this precise, one needs to use tha language of Calculus: if $w(t)$ is the amount of water at time t, then w is a ***continuous function*** of t. The trouble with this is: at this point you only have a nonrigorous, not very precise idea of what a "continuous function" is. You will learn to define the notion of "continuous function", and work with it, and prove things about it, in your next "Advanced Calculus" or "Real Analysis" course.

- And, instead of 7, you can use any denominator you want, and get lengths such as $\frac{5}{2}$, $\frac{12}{5}$, $\frac{29}{17}$, $\frac{236{,}907}{189{,}276}$, and so on.
- Hence, if n and m are any natural numbers, then we can (at least in principle) construct segments of length $\frac{m}{n}$. That is, we can construct segments of length f, for any fraction f.

The measuring numbers such as $\frac{5}{2}$, $\frac{12}{5}$, $\frac{29}{17}$, or $\frac{236{,}907}{189{,}276}$, that can obtained by dividing a counting number m into n equal parts, where n is another counting number, are called ***fractions***.

And this suggests an idea:

Idea 1: *Perhaps the measuring numbers are exactly the same as the fractions.*

In other words: suppose we use the length u of some straight-line segment U as the unit for measuring length. (That is, we call the lenght of this segment "meter", or "yard", or "foot", or "mile", and then we try to express every length in meters, or yards, or feet, or miles.) When we do that, we will of course need fractions to expres some lenghts because, for example, if we measure distances in miles, not every distance will be 1 mile, or 2 miles, or n miles for some counting number n. Some distances will be, say, half a mile, or three quarters of a mile, on thirteen hundredths of a mile, or forty-seven thousandths of a mile[20].

Then Idea 1 suggests that the length of every segment V should be equal to a fraction $\frac{m}{n}$ times u (wnere m, n are natural numbers, i.e., counting numbers). That means that if we divide the segment U into n equal segments of length $w = \frac{u}{n}$, then the length of U is n times w, and the length of V is m times w. So U and V are commensurable. Since we can take U and V to be any two segments we want, we find that ***If Idea 1 i true, then any two segments are commensurable.***

[20]Here is another important difference between counting and measuring numbers: to count things using counting numbers you do not need units, but to measure amounts using measuring numbers you do. If you are asked how many pills there are in a bottle, then you answer "six", or "twenty-five', or whatever, and nobody is going to ask "six what?". But if you are asked how much water there are in the bottle, and you answer "six", then somebody is going to ask "six what?", expecting that you will say something like "six ounces", or "six liters", because if you do not specify the units of your measurement the number you gave is meaningless.

COMMENSURABLE LENGTHS

"Commensurable" means "measurable together". Precisely:

Definition 12.

- *Two segments U, V, are commensurable if you can use a ruler of the same length w to "measure u and v together", that is, to express both lengths u and v as integer multiples mw, nw of the unit of length w.*
- *Two segments U, V, are incommensurable if they are not commensurable.*

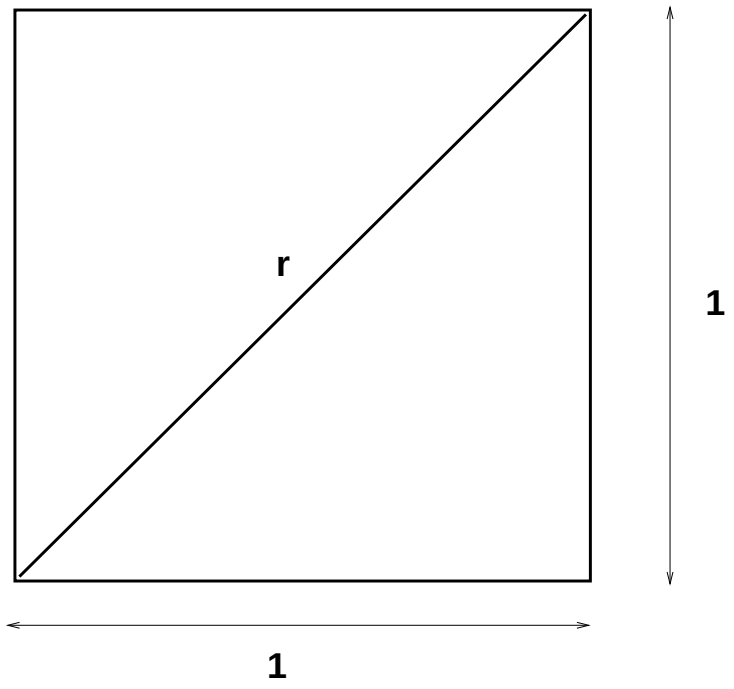
But then a momentous discovery of far-reaching consequences was made:

There are incommensurable lenghts.

That is, ***it is not true that any two lengths are commensurable***.

Precisely: it is possible to construct geometrically[21] a segment whose length r satisfies $r^2 = 2$. For example, if we draw a square whose sides have length 1, then the length r of the diagonal of the square will satisfy $r^2 = 2$, by Pythagoras' theorem.

[21] What does "constructing geometrically" mean? This is tricky. For Euclid (who lived about 23 centuries ago), "constructing geometrically" meant "constructing with a ruler and compass". (See the Wikipedia article "Compass and straightedge consrtuctions".) Using ruler and compass, one can construct lines and circles, but there are lots of other curves—for example, ellipses—that cannot be constructed that way. On the other hand, there are other equally "geometric" methods that can be used to construct some of those curves. For example, ellipses can be constructed using pins and strings. (See the Wikipedia article "Ellipses".)



And it was discovered that ***there is no fraction*** r ***such that*** $r^2 = 2$. This means that

I. If you believe that "number" means "fraction", then there is no number that measures the length of the diagonal of a square whose sides have lengt 1.

II. If you are willing to accept that there could be "numbers" that are not fractions, then maybe there is a number r that measures the length of the diagonal of a square whose sides have lengt 1, but that number r, that we could call "$\sqrt{2}$", is not a fraction.

Today we would say that

- Those numbers that are not fractions, such as $\sqrt{2}$, do indeed exist, and we call them "real numbers".
- The fractions, called "rational[22] numbers", are real numbers, but many

[22]The word "rational" here has nothing to do with "rationality" in the sense of "in

real numbers are "irrational" numbers, that is, numbers that are not rational.

- Actually, most[23] real numbers are not rational.
- It took mathematicians more than 2,000 years after the discovery of the irrationaly of $\sqrt{2}$ to come up with a truly rigorous definition of the concept of "real number". (The name "real number" was introduced by Descartes in the 17th century. The first rigorous definition was given by George Cantor in 1871, and the most widely used definitions were proposed by Karl Weierstrass and Richard Dedekind.

4.4.2 Why was the irrationality of $\sqrt{2}$ so important?

The discovery of the inconmensurability of $\sqrt{2}$ was made, according to legend, by ***Hippasus of Metapontum***, who lived in the 5th century B.C.E and was a member of the religious sect of the Pythagoreans, i.e., the followers of the philosopher and mathematician Pythagoras[24]. And the legend also says that the discovery was so shocking to the Pythagoreans that Hippasus was drowned at sea, as punishment for having divulged the secret. (But this is a legend, and there is no evidence that it is true.)

Why was the existence of inconmensurable magnitudes so upsetting to the Pythagoreans? The reason is this: the Pythagoreans were a mystical-religious cult.

---

accordance withb reason or logic". It comes from the word "ratio", which means "quotient". An "irrational number" is a number that is not the quotient ("'ratio") of two integers. If you hear somebody say something like "scientists have shown that nature is irrational: mathematicians have shown that irrationality is everywhere present, because most numbers are irrational", then you shoud realize that thit is an ignorant statement by somebody who does not understand what "irratioanl numbers" are. The "irrationality" of irrational numbers has nothing to do with their being unreasonable, absurd, or illogical; it just means that they are not quotients of two integers.

[23] If this statement does not strike you as incomprehensible because you don't know what it means, you should think again, and ask yourself "what could it possibly mean to say that most real numbers are irrational"? It turns out that this can be made precise, but making it precise is hard.

[24] Yes, that's the same Pythagoras of Pythagoras's theorem.

> The Pythagoreans honored the effort put into mathematics, and coordinated it with the observation of the cosmos in various ways, for example: by including number in their reasoning from the revolutions and their difference between them, by theorizing what is possible and impossible in the organization of the cosmos from what is mathematically possible and impossible, by conceiving the heavenly cycles according to commensurate numbers with a cause, and by determining measures of the heaven according to certain mathematical ratios, as well as putting together the natural science which is predictive on the basis of mathematics, and putting the mathematical objects before the other observable objects in the cosmos, as their principles.
>
> From the *Wikipedia* article on *Pythagoreanism*, which quotes the *Protrepticus*, by D. S. Hutchinson and M. R. Johnson, a 2015 reconstruction of a lost dialogue of Aristotle.

In other words, for the Pythagoreans everything in the world was determined by ratios (i.e. quotients) of "numbers", and for them "number" meant "natural number" (i.e., counting number). The discovery that some lengths were not ratios of "numbers" undermined the Pythagorean system to such an extent that the members of the sect felt it necessary to conceal this fact from the general public.

But it is important to put all this in proper perspective: there is no real proof that Hippasus truly was the discoverer of the irrationality of $\sqrt{2}$, or that he was drowned at sea for that discovery.

4.4.3 What is a "real number", really?

The discovery that there are lengths that are inconmensurable with one another naturally forced mathematicians to ask a fundamental question: ***what is a "number", really?***

And, as we have explained, it took more than 2,000 years until mathematicians found a satisfactory answer.

4.4.4 The most important number systems: real numbers vs. integers and natural numbers; definition of "rational number"

Now let us look at the main number systems[25] that mathematicians use today.

1. The measuring numbers, together with their negatives, and zero, are called ***real numbers***.
2. The set of all real numbers is called $\mathbb{R}$. (It is also called "the set of all real numbers", or "the real line".)
3. The counting numbers are called ***natural numbers***. (They are also called "positive integers".)
4. The set of all natural numbers is called $\mathbb{N}$.
5. The natural numbers, together with their negatives and zero, are called ***integers***.
6. The set of all integers called $\mathbb{Z}$.
7. The real numbers that are quotients of two integers are called ***rational numbers***. That is, we have the following key definition:

[25] There are many number systems. What we will do here is barely scratch the surface of a very rich theory.

Definition 13.

- *A rational number is a real number r such that there exist integers m, n for which:*

  (a) $n \neq 0$

  (b) $r = \frac{m}{n}$.

- *The set of all rational numbers is called* $\mathbb{Q}$. *(So "$x \in \mathbb{Q}$" is a way of saying "x is a rational number".)*

- *In formal language: If* $r \in \mathbb{R}$, *then* $r \in \mathbb{Q}$ *if*[a]

$$(\exists m \in \mathbb{Z})(\exists n \in \mathbb{Z})\left(n \neq 0 \text{ and } r = \frac{m}{n}\right). \tag{4.27}$$

- *An irrational number is a real number r which is not rational.*

[a] Formula (4.27) is not yet completely formal, because it contains the word "and". Soon we are going to leanr the symbol "$\wedge$" for "and", and then we will be able to rewrite (4.27) as $(\exists m \in \mathbb{Z})(\exists n \in \mathbb{Z})\left(n \neq 0 \wedge r = \frac{m}{n}\right)$.

4.4.5 A remark about sets

We will spend a lot of time in this course studying ***sets***. At this point, all you need to know is that

- ***sets have members***.
- If S is a set and x is an object (for example, a number or a person or a giraffe or a set) then "$x \in S$" is a way of saying that x is a member of S.
- "$x \in S$" is read as "x belongs to S", or "x is in S", or "x is a member of S".
- We write "$x \notin S$" to indicate that x is not a member of S.
- So, for example,

  - If C is the set of all cows, then to say that Suzy is a cow we can equally well say "Suzy$\in C$".
  - You can read "Suzy$\in C$ in any of the following ways:
    1. Suzy belongs to C,
    2. Suzy is in C,
    3. Suzy belongs to the set of all cows,
    4. Suzy is a cow.

    But the third reading, although correct, is very stupid, because there is no reason to say "Suzy is a member of the set of all cows" when you can say the same thing in a much shorter and simpler way by saying "Suzy is a cow".
  - Similarly, you can read "Suzy$\notin C$ in any of the following ways:
    1. Suzy does not belong to C,
    2. Suzy is not in C,
    3. Suzy does not belong to the set of all cows,
    4. Suzy is not a cow.

    And the third reading, though correct, sounds silly, so you would never say it that way.

- Here is another example.
  - "$\mathbb{N}$", as we know, is the set of all natural numbers. So, to say that 3 is a natural number we can equally well say "$3 \in \mathbb{N}$".
  - You can read "$3 \in \mathbb{N}$ in any of the following ways:
    1. 3 belongs to $\mathbb{N}$,
    2. 3 is in $\mathbb{N}$,
    3. 3 belongs to the set of all natural numbers,
    4. 3 is a natural number.

    But the third reading, although correct, is very stupid, because there is no reason to say "3 is a member of the set of all natural number" when you can say the same thing in a much shorter and simpler way by saying "3 is a natural number".

Problem 25. For each of the following formulas,

(a) translate the formula into English,
(b) indicate whether it is true or false.

Give the best, most natural English translation. For example, the formula "$1 \in \mathbb{N}$" could be translated as "1 belongs to the set of natural numbers", but this sounds very awkward. A much better way to say the same thing in English is "1 is a natural number", so this translation is to be preferred.

1. $-3 \in \mathbb{N}$,
2. $0 \in \mathbb{N}$,
3. $0 \notin \mathbb{Z}$,
4. $0 \in \mathbb{Z}$,
5. $-3 \in \mathbb{R}$,
6. $0 \in \mathbb{R}$,
7. $0 \notin \mathbb{R}$,
8. $0 \in \mathbb{R}$,
9. $0 \in \mathbb{Q}$,
10. $3 \in \mathbb{Q}$,
11. $-3 \in \mathbb{Q}$,
12. $\frac{237}{42} \in \mathbb{Q}$,
13. $\sqrt{2} \in \mathbb{Q}$,
14. $\sqrt{2} \notin \mathbb{Q}$,
15. $\pi \in \mathbb{Q}$.

4.5 The irrationality of $\sqrt{2}$

As explained before, we could state the theorem on the irrationality of $\sqrt{2}$ by saying that "$\sqrt{2}$ is irrational". This, however, would mean that there is a "number $\sqrt{2}$", i.e., a number whose square is 2. But the issue whether such a number exists is different from the one that concerns us here, namely, whether there exists a rational number r such that $r^2 = 2$. So I prefer to state the theorem in a way that does not imply any *a priori* commitment to the existence of a "number" r such that $r^2 = 2$.

And, before we give the proof, we introduce a few concepts and state some facts that will be used in the proof, (These facts will be proved later in the course.)

4.5.1 Even and odd integers

THE DEFINITION OF "EVEN" AND "ODD" INTEGERS

Definition 14. *Let a be an integer. We say that a is <u>even</u> if it is divisible by 2. And we say that a is <u>odd</u> if it is not even.*

4.5.2 Coprime integers

The integers 1 are -1 are factors of every integer, because if $n \in \mathbb{Z}$ then $n = n \times 1$ and $n = (-n) \times (-1)$, so n is divisible by 1 and by -1. So 1 and -1 are not very interesting factors, because they are always there. So we refer to 1 and -1 as the ***trivial factors*** of an integer.

THE DEFINITION OF "COPRIME INTEGERS"

Definition 15.

- *Let a, b be integers. We say that a and b are coprime if they do not have any nontrivial common factors.*
- *We write "$a \perp b$" to indicate that a and b are coprime.*
- *In formal language, if $a \in \mathbb{Z}$ and $b \in \mathbb{Z}$, then $a \perp b$ if*

$$\sim (\exists k \in \mathbb{Z})(k|a \text{ and } k|b \text{ and } k \neq 1 \text{ and } k \neq -1)\,.$$

Example 9. The integers 12 and 35 are coprime. Indeed:

- The factors of 12 are 1, -1, 2, -2, 3, -3, 4, -4, 6, -6, 12 and -12.
- The factors of 35 are 1, -1, 5, -5, 7, -7, 35 and -35.

So the only common factors are 1 and -1, i.e., the trivial factors. Hence 12 and 35 are coprime. □

4.5.3 Proof of the irrationality of $\sqrt{2}$

Now, finally, we are ready to prove that $\sqrt{2}$ is irrational.

We are going to use two facts:

Fact 1. *Every rational number is equal to a quotient $\frac{m}{n}$ of two coprime integers.*

Fact 2. *The product of two odd integers is odd.*

Example 10. Here are some examples to illustrate what Fact 1 means:

- let $a = \frac{-36}{22}$. The integers -36 and 22 are not coprime, because they are both divisible by 2. But we can factor out the 2, and get $a = \frac{-18}{11}$. Now the numerator -18 and the denominator 11 are coprime.

- let $a = \frac{630}{840}$. The natural numbers 630 and 840 are not coprime, because they are both divisible, for example, by 2. We can factor out the 2, and get $a = \frac{315}{420}$. The numerator 315 and the denominator 420 are not yet coprime, because they are both divisible, for example, by 3. We can factor out the 3, and get $a = \frac{105}{140}$. The numerator 105 and the denominator 140 are not yet coprime, because they are both divisible, for example, by 5. We can factor out the 5, and get $a = \frac{21}{28}$. The numerator 21 and the denominator 28 are not yet coprime, because they are both divisible by 7. We can factor the common factor 7 and we get, finally, $a = \frac{3}{4}$. And now the numerator 3 and the denominator 4 *are* coprime. □

Theorem 9. *There does not exist a rational number r such that $r^2 = 2$.*

Proof. We give a proof by contradiction .

Assume that there exists a rational number r such that $r^2 = 2$.

Pick one such number and call it r. (Here we are using Rule $\exists_{use}$.)

Using the fact that $r \in \mathbb{Q}$, we may pick integers m, n such that

(1) $n \neq 0$,
(2) $r = \frac{m}{n}$,

(Here we are using again Rule $\exists_{use}$.)

Using Fact 1, we may actually choose m, n such that

(3) m and n are coprime.

Since $r^2 = 2$, we have $\frac{m^2}{n^2} = 2$.

Therefore $m^2 = 2n^2$.

So m^2 is even.

But then m is even. (Reason: Assume[26] that m is not even. Then m is odd. So by Fact 2, m^2 is odd. But we have proved that m^2 is even. So m^2 is not odd. Therefore m^2 is odd and m^2 is not odd, which is a contradiction.)

Since m is even, m is divisible by 2, that is, $(\exists k \in \mathbb{Z})m = 2k$.

[26]Notice that we have a proof by contradiction within our main proof by contradiction.

So we may pick an integer k such that $m = 2k$.

Then $m^2 = 4k^2$.

But $m^2 = 2n^2$.

Hence $2n^2 = m^2 = (2k)^2 = 4k^2$.

Therefore $n^2 = 2k^2$.

So n^2 is even.

But then n is even. (Reason: Assume[27] that n is not even. Then n is odd. So n^2 is odd by Fact 2. But we have proved that n^2 is even. So n^2 is not odd. Therefore n^2 is odd and n^2 is not odd, which is a contradiction.)

So m is even and n is even.

Therefore m and n are divisible by 2.

So m and n have a nontrivial common factor.

Hence $\boxed{m \text{ and } n \text{ are not coprime}}$.

But $\boxed{m \text{ and } n \text{ are coprime}}$

So $\boxed{m \text{ and } n \text{ are coprime and } m \text{ and } n \text{ are not coprime}}$, which is a contradiction.

So the assumption that there exists a rational number r such that $r^2 = 2$ has led us to a contradiction.

Therefore $\boxed{\text{there does not exist a rational number } r \text{ such that } r^2 = 2}$.**Q.E.D.**

4.6 More irrationality proofs

We now use the same technique to prove that $\sqrt{3}$ is irrational. The key point here is to realize that "even vs. odd" now has to be replaced by "divisible by 3 vs. not divisible by 3". And, in order to do the crucial step (the analogue of "if m^2 is divisible by 2 then m is divisible by 2") we need a generalization of Fact 2:

[27] Another proof by contradiction !

Fact 3. *If p is a prime number, then the product of two integers that are not divisible by p is not divisible by p either.*

(We will prove Fact 3 later.)

Theorem 10. *There does not exist a rational number r such that $r^2 = 3$.*

Proof. We want to prove that $\sim (\exists r \in \mathbb{Q}) r^2 = 3$. We will do a proof by contradiction .

Assume that $(\exists r \in \mathbb{Q}) r^2 = 3$, i.e., there exists a rational number r such that $r^2 = 3$.

Pick one such number and call it r.

Using the fact that $r \in \mathbb{Q}$, we may pick integers m, n such that

(1) $n \neq 0$,
(2) $r = \frac{m}{n}$,

Then, using Fact 1, we can actually choose m, n so that

(3) m and n are coprime.

Since $r^2 = 3$, we have $\frac{m^2}{n^2} = 3$.

Therefore $m^2 = 3n^2$.

So m^2 is divisible by 3.

But then m is divisible by 3. (Reason: By Fact 3, if m was not divisible by 3, it would follow that m^2 is not divisible by 3 either. But m^2 is divisible by 3, and we got a contradicition.)

Since m is divisible by 3, we may pick an integer k such that $m = 3k$.

Then $m^2 = 9k^2$.

But $m^2 = 3n^2$.

Hence $3n^2 = 9k^2$, so

$$n^2 = 3k^2 \,. \tag{4.28}$$

So n^2 is divisible by 3.

But then n is divisible by 3. (Reason: By Fact 3, if n was not divisible by 3, it would follow that n^2 is not divisible by 3 either. But n^2 is divisible by 3, and we got a contradicition.)

So 3 is a factor of m and 3 is a factor of n.

Hence m and n have a nontrivial common factor.

So m and n are not coprime.

But m and n are coprime.

Therefore $\boxed{m \text{ and } n \text{ are coprime and } m \text{ and } n \text{ are not coprime}}$, which is a contradiction,

So the assumption that there exists a rational number r such that $r^2 = 3$ has led us to a contradiction,

Therefore $\boxed{\text{there does not exist a rational number } r \text{ such that } r^2 = 3}$.**Q.E.D.**

4.6.1 What happens when you make a mistake in a proof

Can we do the same that we did before to prove the following theorem?

THEOREM: There does not exist a rational number r such that $r^2 = 4$.
Proof. We will do a proof by contradiction .

Assume that there exists a rational number r such that $r^2 = 4$.

Pick one such number and call it r.

Using Fact 1, we may pick integers m, n such that

(1) $n \neq 0$,
(2) $r = \frac{m}{n}$,
(3) m and n have no nontrivial common factors.

Since $r^2 = 4$, we have $\frac{m^2}{n^2} = 4$.

Therefore $m^2 = 4n^2$.

So m^2 is divisible by 4.

But then m is divisible by 4. (Reason: By Fact 3, if m was not divisible by 4, it would follow that m^2 is not divisible by 4 either. But m^2 is divisible by 4, and we got a contradicition.)

Since m is divisible by 4, we may pick an integer k such that $m = 4k$.

Then $m^2 = 16k^2$.

But $m^2 = 4n^2$.

Hence $n^2 = 4k^2$, so

$$n^2 = 3k^2 \,. \tag{4.29}$$

So n^2 is divisible by 4.

But then n is divisible by 4. (Reason: By Fact 3, if n was not divisible by 4, it would follow that n^2 is not divisible by 3 either. But n^2 is divisible by 4, and we got a contradicition.)

So 3 is a factor of m and 4 is a factor of n.

Hence m and n have a nontrivial common factor.

So m and n are not coprime.

But m and n are coprime.

Therefore $\boxed{m \text{ and } n \text{ are coprime and } m \text{ and } n \text{ are not coprime}}$, which is a contradiction,

So the assumption that there exists a rational number r such that $r^2 = 4$ has led us to a contradiction,

Therefore $\boxed{\text{there does not exist a rational number } r \text{ such that } r^2 = 4}$.**Q.E.D.**

Same proof, right?

WRONG!!!!!

What is wrong here?

1. The result is ***false***. It is not true that there does not exist a rational number r such that $r^2 = 4$. Indeed, if we take $r = 2$ then r is ratinal and $r^2 = 4$.

2. Since the conclusion of the proof is false, the proof itself must be wrong. That is, whoever wrote this proof must have cheated[28] in some step.

[28] Nothing personal here. "Cheat" means "violate the rules." Of course, I haven't told you yet what the rules are, but let me anticipate one of them. ***You are allowed to use a result that has been proved, but you are now allowed to make up a statement that has not been proved and use it as if it was true.***

In our case, Fact 3 explicitly says that "if p is prime then if a is not divisible by p it follows that a^2 is not divisible by p". So we are allowed to apply Fact 3 if p is prime, but we are not allowed to apply it if p is not prime.

So the two steps where we applied Fact 3 are wrong. In those steps, we cheated, by violating the rules.

The general principle is this: ***If a proof is correct then you can be sure that the conclusion is true.***

And another way to say that is this: ***if the conclusion of a proof is false, then the proof must be wrong. There has to be a mistake in the proof itself.***

So, if I give you a proof of a conclusion that is false, you have to be able to find where in the proof the author cheated. I will not be satisfied with a statement such as "the proof is wrong because the conclusion is false." I will want to know where in the proof a mistake was made.

Consider the following analogy: If I am trying to drive to Boston and end up in New York, then of course I can conclude thta I did something worng. But I will want to know what I did wrong, where I made a wrong turn. The same happens with proofs.

4.6.2 More complicated irrationality proofs

I hope it is clear to you that the same method, exactly, will apply to prove that $\sqrt{5}$, $\sqrt{7}$, $\sqrt{11}$, and, more generally, $\sqrt{p}$ for any prime number, is irrational.

Now let us try a more complicated case. Let us prove that

Theorem 11. *There does not exist a rational number r such that $r^2 = 12$.*

Remark 5. The number 12 is not prime. (Actually, $12 = 4 \times 3$.) So we cannot apply Fact 3 with 12 in the role of p.

Proof. We will do a proof by contradiction .

Assume that there exists a rational number r such that $r^2 = 12$.

Pick one such number and call it r, so $r^2 = 12$..

Using the fact that $r \in \mathbb{Q}$, we may pick integers m, n such that

(1) $n \neq 0$,

(2) $r = \frac{m}{n}$,

Then, using Fact 1, we may pick m, n such that

(3) m and n are coprime.

Since $r^2 = 12$, we have $\frac{m^2}{n^2} = 12$.

Therefore $m^2 = 12n^2$.

Hence $m^2 = 3 \times 4n^2$.

So m^2 is divisible by 3.

But then m is divisible by 3. (Reason: By Fact 3, if m was not divisible by 3, it would follow that m^2 is not divisible by 3 either. But m^2 is divisible by 3, and we got a contradicition.)

Since m is divisible by 3, we may pick an integer k such that $m = 3k$.

Then $m^2 = 9k^2$.

But $m^2 = 12n^2$.

Hence $12n^2 = 9k^2$, so

$$4n^2 = 3k^2 \,. \tag{4.30}$$

So $4n^2$ is divisible by 3.

But then n is divisible by 3. (Reason: By Fact 3, assume n is not divisible by 3; then by Fact 3 n^2 is not divisible by 3; since 4 is not divisible by 3, another application of Fact 3 tells us that $4n^2$ is not divisible by 3. But $4n^2$ is divisible by 3, so we got a contradiction.)

So 3 is a factor of m and 3 is a factor of n.

Hence m and n have a nontrivial common factor.

So m and n are not coprime.

But m and n are coprime.

Therefore $\boxed{m \text{ and } n \text{ are coprime and } m \text{ and } n \text{ are not coprime}}$, which is a contradiction,

So the assumption that there exists a rational number r such that $r^2 = 12$ has led us to a contradiction,

Therefore $\boxed{\text{there does not exist a rational number } r \text{ such that } r^2 = 12}$.**Q**.**E**.**D**.

Problem 26. ***Prove*** that each of the following numbers is irrational:

1. $\sqrt{5}$,
2. $\sqrt[3]{5}$,
3. $\sqrt[3]{9}$,
4. $\sqrt{28}$,
5. $\sqrt{2+\sqrt{2}}$,
6. $\sqrt{\frac{2}{3}}$,
7. $\sqrt{\frac{27}{31}}$. □

Problem 27. ***Prove or disprove***[29] each of the following statements:

1. The sum of two rational numbers is a rational number.
2. The product of two rational numbers is a rational number.
3. The sum of two irrational numbers is an irrational number.
4. The product of two irrational numbers is an irrational number.
5. The sum of two irrational numbers is a rational number.
6. The product of two irrational numbers is a rational number.
7. The sum of a rational number and an irrational number is an irrational number.
8. The product of a rational number and an irrational number is an irrational number. □

Problem 28.

I. ***Explain*** why the following "proofs" that $\sqrt{2}+\sqrt{3}$ and $\sqrt{6}$ are irrational (in which we are allowed to use the facts that $\sqrt{2}$ and $\sqrt{3}$ are irrational) are wrong:

1. *Proof that $\sqrt{2}+\sqrt{3}$ is irrational:*

   We know that $\sqrt{2}$ is irrational.
   We know that $\sqrt{3}$ is irrational.

[29] To ***disprove*** a statement means "to prove that the statement is false". For example, when we proved that 1 is not even we disproved the statement '1 is even".

Hence the sum $\sqrt{2}+\sqrt{3}$ is irrational. **Q**.**E**.**D**.

2. *Proof that* $\sqrt{6}$ *is irrational:*

We know that $\sqrt{2}$ is irrational.
We know that $\sqrt{3}$ is irrational.
Hence the product $\sqrt{2}.\sqrt{3}$ is irrational.
So $\sqrt{6}$ is irrational. **Q**.**E**.**D**.

II. ***Give correct proofs*** that $\sqrt{2}+\sqrt{3}$ and $\sqrt{6}$ are irrational. □

Problem 29. ***Prove*** that $\sqrt{2}+\sqrt[3]{2}$ is irrational. □

Problem 30. ***Prove*** that $\sqrt{2}+\sqrt{3}+\sqrt{5}$ is irrational. (NOTE: This requires some hard thinking on your part.) □

Problem 31. ***Prove*** that $\sqrt{2}+\sqrt{3}+\sqrt{5}+\sqrt{7}$ is irrational. (NOTE: This requires ***quite a lot*** of thinking on your part.) □

Problem 32. ***Prove*** that, if $n \in \mathbb{N}$, and $p_1, p_2, \ldots, p_n$ are n distinct primes, then $\sqrt{p_1}+\sqrt{p_2}+\cdots+\sqrt{p_n}$ is irrational. (NOTE: This is very difficult.) □

5 What is a proof, really?

THIS SECTION IS STILL BEING WRITTEN. WHEN IT IS FINISHED IT WILL BE INCLUDED IN THESE NOTES.

5.1 Analysis of the proof of Theorem 5

THIS SECTION IS STILL BEING WRITTEN. WHEN IT IS FINISHED IT WILL BE INCLUDED IN THESE NOTES.

6 The languages of mathematics: formal, natural, and semiformal

In these notes, we will be talking mostly about ***mathematical objects***, that is, numbers of various kinds (natural numbers, integers, rational numbers, real numbers, complex numbers, integers modulo n, etc.), sets, functions, relations, graphs, geometric objects (such as points, lines, segments, angles, circles, planes, curves and surfaces of various kinds, etc.), and many other kinds of objects (such as groups, rings, fields, algebras, modules, vector spaces, manifolds, bundles, Lie groups, etc.) that mathematicians have invented and you will learn about in more advanced courses.

And we will talk about these mathematical objects using ***mathematical language***. But mathematical language is a special kind of language, in many ways similar to other languages such as English, and in many ways different. So, in order to talk about mathematical language we will want to say a few words about language in general, so that we can explain what makes mathematical language special.

Mathematical language, as commonly used, is ***semiformal language***, that is, a mixture of ***formal language*** and the ***natural language*** (English, Chinese, French, whatever) that one uses in a particular country. (Formal lamguage is a language consisting entirely of formulas. For example, the statement "$A = \pi R^2$" is an expression in formal language.)

For example, when we say

$$\text{from the facts that } 2+2=4 \text{ and } 4+2=6 \text{ we deduce that } (2+2)+2=6 \tag{6.31}$$

this is a mixture of formal mathematical language and English. (The formal language part consists of the formulas "$2+2=4$", "$4+2=6$", and "$4+2=6$". The English part is the rest.)

If we wanted to say the same thing in French, we would say

$$\text{des faits que } 2+2=4 \text{ et } 4+2=6 \text{ on deduit que } (2+2)+2=6\,. \tag{6.32}$$

Notice that ***the formal language part does not change***. That's because ***formal language is universal***. The formula "$2+2=4$" is exactly the same in English, French, Chinese, or any other language.

As we will see in the course, ***it is possible to formalize mathematics fully***, that is, to develop a formal language into which we can translate every mathematical statement.

For example, statement (6.31) would become, in purely formal language:

$$(2+2=4 \wedge 4+2=6) \Longrightarrow (2+2)+2=6\,. \tag{6.33}$$

And, once you get to this level, the texts you get are no longer in English or Franch or Chinese, because ***formal language is the same everywhere***, exactly as the formula "$1 + 1 = 2$' is the same everywhere and can be understood by all people, no matter what language they speak.

This means that if we could write all of mathematics in formal language, we would have a language that permits people of all nationalities, speaking all kinds of lamguages, to communicate easily: if a mathematician who speaks Chinese says something, and a mathematician who speaks English does not understand, then all these two mathematicians have to do is switch to formal language, and then they would have no problem communicating.

Formal language has other advantages that we will talk about soon. So you would think that mathematicians must use formal language all the time. But in fact we do not. We use a semiformal language which is a mixture of formal language and our own natural languages, because formal language is too dry and to hard to read. But formal language remains the means of communication of last resort: if I don't understand something you wrote, then I would ask you to say it in formal language. I you cannot say it in formal language, then what you wrote is meaningless. If you can say it in formal language, then I will understand what you said, and I will be able to decide if it is right or wrong.

Example 11. Suppose you are trying to define "prime number", and write "a prime number is a number that is only divisible by 1 and itself". Then I do not understand what you are saying, so I cannot tell if it is right or wrong.

Why do I not understand?

- Fisrt of all, I do not understand what "number" means. There are lots of different kinds of numbers: natural numbers, integers, rational numbers, real numbers, complex numbers, integers modulo n, etc. When you say "number", which one do you mean?

- Also: what does "only divisible" mean? You may say that when yoy write "p is only divisible by 1 and itself", what you mean is that "the only factors of p are 1 and p". But then I would reply: "so 3 is not prime, because the factors of 3 are 3, 1, -1 and -3, so it's not true that the only factors are 1 and 3; so 3 is not prime." Then you would probably reply: "I did not mean to count negative factors as factors", And I would aswer: "why didn't you say that?"

If I ask you to write your statement in formal language, then that will force you to make your meanings precise. For example, you will write something

like[30]

$$\text{if } p \in \mathbb{N}\,, \text{then } p \text{ is } \underline{\text{prime}} \text{ if } (\forall k \in \mathbb{N})\Big(k|p \Longrightarrow (k = 1 \vee k = p)\Big)\,. \quad (6.34)$$

This is now completely clear, so at this point I will finally have understood what you are saying. And then I will be able to tell if this is right or wrong.

The answer is: as a definition of "prime number", this is wrong, because 1 is not prime, but according to (6.34) 1 is prime.

But we can make it right by writing:

$$\text{if } p \in \mathbb{N}\,, \text{then } p \text{ is } \underline{\text{prime}} \text{ if } p > 1 \wedge (\forall k \in \mathbb{N})\Big(k|p \Longrightarrow (k = 1 \vee k = p)\Big)\,. \quad (6.35)$$

6.1 Things and their names

In any language, whether it is English, French, Russian, Spanish, Chinese, or formal or semiformal mathematical language, we talk about ***things*** (objects, entities), and in order to do that we give them ***names***.

[30] This is not yet a fully formal definition. To make it fully formal we need to introduce a symbolic way to say "p is prime". We can do this by using "$P(x)$" for "x is prime", and then your statement would become: $(\forall p \in \mathbb{N})\Bigg(P(p) \Longleftrightarrow (\forall k \in \mathbb{N})\Big(k|p \Longrightarrow (k = 1 \vee k = p)\Big)\Bigg)$. This is not yet a correct definition of "prime number" but at least it is pefectly clear.

THINGS

In these notes, the word ***thing*** refers to an object of any kind: a concrete inanimate material object such as a table or a molecule or a planet, a "living thing" such as a plant, an animal, a person, or an amoeba, or an abstract thing such as a mathematical object.
So, in these notes, Mount Everest is a thing, and the chair on which you are sitting is a thing, and a book is a thing, but so are a giraffe, a spider, and you, and I, and my uncle Jim, and the number four, and the set $\mathbb{N}$ of all natural numbers.
Some students don't like using the word "thing" to refer to people, perhaps because they are thinking that "people are not things". My answers to that are:

1. We can use words in any way we like, as long as we do it consistently. So in this course we can decide how to use the word "thing", and there should be no problem as long as what we mean is clear to everybody.

2. We often do talk about "living things", and that includes people.

3. If you don't like using the word "thing" in this way, there is a word that's perfect for you: you can talk about "entities" instead. An entity is anything that exists. It can be a table, a river, a planet, an atom, a cell, a plant, a giraffe, a person, a number, a triangle, a matrix, a set, or a function. So just substitute the word "entity" for "thing" throughout these notes, and you will be fine.

6.1.1 Giving things individual names

The simplest way to give names to things is to give each thing an individual name, as when you call people with names such as "Mary", "John", or "George Washington", you give cities names such as "New York City", "Paris", or "London", or you give mountains names such as "Mount Everest" or "Mount Aconcagua".

But this way of naming things is not very convenient, because in our daily life we have to talk about an enormous number of things of many different kinds, and it would be truly impossible to give an individual name to each one.

Just imagine if every fork, every knife, every spoon, every plate, every

glass, every cup, every napkin, every table, every pencil, every pen, every cell phone, every toothbrush, every animal, every plant, every cell in every person's or animal's or plant's body, every molecule and every atom in the Universe, every electron and every proton and every neutron and every particle of every kind, had to have its own individual name, and you had to know the name of each of those things before you can talk about it! Imagine how difficult life would be if every time you want to ask a waiter for a spoon you had to find out first the name of that particular spoon!

6.1.2 Variable noun phrases

So languages have developed a special device for naming things without having to give each individual thing its own name. We do this by using ***variables***, that is, noun phrases that can be temporarily designated to stand for a particular thing but can then be ***re-used***, as needed, to stand for a different thing.

NOUN PHRASES

A ***noun phrase*** is a word or phrase that stands for or is the name of something or somebody. For example: "he", "she", "the giraffe", "my uncle Jimmy', "Mount Everest", "the pencil", "the Math 300 final exam", "the table that I bought yesterday", "the President of the United States", "Mary", "New York City", "the most expensive restaurant in New York City", "the owner of the most expensive restaurant in New York City", are all noun phrases.

Example 12. When I say "I am going to open the door and let you in", the noun phrases "I", "the door", and "you" stand, respectively, for the speaker, a door, and the person that the speaker is talking to. But later, if somebody else says the same thing to somebody else, the words "I", "the door", and "you" will stand for two different people and a different door.

These noun phrases are ***variables***: at each particular time they are used they stand for some definite thing or person, called the ***referent***, or the ***value*** of the variable. In each particular instance, it must be clear what the value is. (For example, if you and I are on a beach, and there is no door in sight, then when I say "I am going to open the door and let you in" you will not understand what I am talking about[31].). □

[31] Unless my statement is part of some larger context that makes the value of the noun

Variable noun phrases are re-usable: after I have used "the door" to refer to one particular door, I may use "the door" again later to refer to a different door.

Example 13. In a court of law, the noun phrase "the defendant" is used as a variable. When a trial begins, someone announces in some way that, for the duration of this trial, the words "the defendant" will refer to a certain specific person. Then, during the trial, everybody refers to that person as "the defendant". When the trial is over, the variable "the defendant" becomes ***free***, that is, not attached to any particular person, and is free to be used to refer to a new defendant when a new trial begins. □

Example 14. When you buy a house, the contract will probably contain a clause at the beginning declaring the words "the buyer" to stand for you for that particular contract. This means that the phrase "the buyer" is a variable, whose value is you for this contract. Later, for a new house sale, where the buyer is a different person, a new contract will be signed, in which the phrase "the buyer" has a totally different value. So the value of the phrase "the buyer" is fixed only within a specific contract, and changes when you go to another contract. □

6.1.3 Declaring the value of a variable

When we communicate our thoughts by speaking or writing, we use variable noun phrases all the time. But in order to be understood we also have to communicate to the reader or listener what each variable stands for each time we use it. That is, we have to ***declare*** the values of the variables we use. How is that done?

In English, values of variables are declared in dozens of different ways. For example,

- Often, we first mention a person by his or her name, and then when we use the pronouns "he", "him", "his", "she", "her", it is understood that the pronoun stands for that person. For example, suppose I write

  > George Washington was the first president of the United States, and ***he*** served as president for two terms. ***He*** was succeeded by John Adams, who served only one term. When Adams ran for reelection to a second term, ***he*** was the object of malicious attacks by his opponents, and eventually lost the election to Thomas Jefferson.

phrase "the door" clear. For example, I could be telling you that later, when we get home, I will open the door and let you in. In that context, the value of "the door" is clear.

In this text, the pronoun "he" appears three times. The first two times, it clearly refers to George Washington, but the third time it refers to John Adams. The mention of John Adams undoes the declaration that "he" stands for George Washington, and assigns the new value "John Adams" to the pronoun.

- The pronoun "I" is understood to stand for whoever is speaking or writing.
- The pronoun "you" is understood to stand for whoever the speakers or writers are addressing themselves to.
- Values of variables are often declared by pointing. For example, if I say "please give me that book", and I point to a book, then that book is the value of the variable "the book".
- Sometimes, the value of a variable is clearly determined by the fact that there is only one thing within sight that the variable can stand for. For example, if I say "please give me the book", and there is only one book within sight, then that book is the value.
- Often, the value of a variable is announced explicitly, as in the examples we gave above of the variable "the defendant" in a trial, and "the buyer" in a contract.

6.1.4 Using variables to name things in mathematical language

In mathematical language, it is customary to use ***letters*** as variables. The most commonly used letters are

- lower case letters such as x, y, r, p, q, a, b, atc.,
- capital letters such as X, Y, P, Q, A, B, etc.,
- lower case Greek letters (α, β, φ, ψ, σ, etc.),
- capital Greek letters[32] (Φ, Ψ, Σ, etc.).

But it is perfectly possible to use as variables other symbols such as

- longer strings such as "*abb*" or "the number I have been talking about",

[32]Some capital Greek letters are not used, because they are identical to their Latin counterparts. For example, A (capital alpha) and B (capital beta) are identical to the Latin A and B.

- other symbols, such as $\diamond$, or $\clubsuit$.

Actually, ***you can use as a variable any symbol or string of symbols you want*** (except only for symbols such as $=$, $<$, $\leq$, $>$, $\geq$, $+$, $\times$, $\to$, $\Rightarrow$, $\wedge$, $\vee$, $\Leftrightarrow$, etc., that already stand for something else), ***provided that you declare its value*** (i.e. tell the reader clearly what the symbol or string of symbols stands for).

Remark 6. The symbols π and e stand for the well known real numbers $3.141592653589793238\ldots$ and $2.718281828459045235\ldots$, respectively. But even those symbols can be (and sometimes are) used as variables with other values, provided that the reader is told clearly what these symbols stand for[33]. □

6.1.5 Free (i.e. open) vs. bound (i.e. closed) variables

A free variable (or "open variable") in a text is a letter (or string of symbols) that is "unattached", in the sense that it has not been assigned a value, and is therefore free to be assigned any value we want.

A bound variable (or "closed variable") is a variable that has been assigned a value.

For instance, suppose a student starts a proof by writing:

(*) $x^2 = 1 + x\,.$

or

(**) I am going to prove that $x^2 = 1 + x\,.$

In these texts, the letter x is a free variable. The formula says that "x-squared is equal to $x + 1$", but it does not tell us who x is. So we have no way to know whether the formula is true or false. Therefore ***texts such as (*) or (**) are unacceptable, because they are meaningless.***

On the other hand, suppose a student writes

(***) Let $x = \frac{1+\sqrt{5}}{2}$.
Then
$$x^2 = 1 + x\,.$$

[33]For example: the symbol π is sometimes used to stand for a permutation; the expression $\pi_k(S)$ stands for the k-th homotopy group of a space S; the letter e is sometimes used for the charge of an electron.

In this text, ***the phrase "let $x = \frac{1+\sqrt{5}}{2}$" effectively declares the variable x to have the value $\frac{1+\sqrt{5}}{2}$.***

So, after this value declaration, "x" stands for the number $\frac{1+\sqrt{5}}{2}$.

Then the meaning of (***) is perfectly clear, so ***(***) is acceptable, because in it the variable x is used correctly: before it is used, a value for it is declared.***

And then the meaning of (***) is perfectly clear: (***) is just a roundabout way to say that

$$\left(\frac{1+\sqrt{5}}{2}\right)^2 = 1 + \frac{1+\sqrt{5}}{2}\,.$$

Once this particular use of the variable x is over, you could, if you want to, use the same letter to represent some other number or object of any kind. But in that case it would have to be very clear that the old declaration that $x = \frac{1+\sqrt{5}}{2}$ no longer applies.

You could do this, for example, by saying something like

(****)

> Let $x = \frac{1+\sqrt{5}}{2}$. Then $x^2 = 1 + x$.
> Now suppose, instead, that $x = \frac{1-\sqrt{5}}{2}$. Then it is also true that $x^2 = 1 + x$.

In (****), the word "now" serves the purpose of telling the reader that "we are starting all over again, and the old declared value of x no longer applies." (And the word "instead", which is unnecessary, strictly speaking, reinforces that.)

6.1.6 What does "arbitrary" mean

There is another way to assign a value to a variable: we can declare the value to be an ***arbitrary*** object of a certain kind.

ARBITRARY THINGS

An ***arbitrary thing*** of a certain kind is a fixed thing about which we know nothing, except that it is of that kind. For example, an "arbitrary integer" is an integer about which you know nothing other than that it is an integer.

The way you should think about "arbitrary things " is as follows.

- Imagine that you are playing a game against somebody (a friend, or a computer, or an alien from another planet) that we will call the **CAT** ("creator of arbitrary things").
- The CAT's job is as follows: every time you say or write "let a be an arbitrary thing of such and such kind," the CAT picks one such thing, writes down what that thing is on a piece of paper, puts the paper in an envelope, and seals the envelope.
  So, for example, if you say "let a be an arbitrary natural number" then the CAT will pick a natural number and write down what it is on a piece of paper that will go inside the envelope.
- Later. after you have finished talking or writing, you or the CAT will open the envelope, and you will know who a really was.
- At that point,
  - if what you said about a turns out to be true, then you win, and the CAT loses.
  - if what you said about a is not true, then the CAT wins, and you lose.

The key fact is this: ***In order to win, you have to be sure that everything you say about*** a ***is true of all the things of the given kind,*** because if there is just one thing for which what you said is not true, then a could turn out to be that thing, and then you will have been proved wrong, and will lose.

Example 15. Suppose you say:

Let n be an arbitrary integer.

What can you say after that, being sure that it is true?

Certainly, you cannot say that $n = 2$, because n could be 1, or -7, or 25.

And you cannot say that n is even, because n could be odd.

But here are a few things you *can* say:

- $n = n$.
- $|n| \geq 0$.
- n is either a natural number, or the negative of a natural number, or zero.
- $n + n^2$ is even. (Reason: n is either even or odd. If n is even, then n^2 is also even, so the sum $n + n^2$ is even. If n is odd, then n^2 is also odd, and the sum of two odd integers is even, so $n + n^2$ is even. So, no matter who n is, whether it is even, or odd, positive or negative, yuuo can be sure that $n + n^2$ is even.)
- $n^2 \geq 0$. (Reason: the square of every real number, and in particular of every integer, is ≥ 0.)
- If n is even then n^2 is divisible by 4. (This sentence is true for ***every*** natural number n. Indeed, the sentence is an implication: n is even$\Longrightarrow$ n^2 is divisible by 4. The integer n could be even or odd, and you have no control over that, because the CAT chooses n, and the CAT can choose n any way he or she wamts to. But: if n is odd, then the implication "n is even$\Longrightarrow$ n^2 is divisible by 4" is true, because the premise "n is even" is false; and if n even then we may pick an integer k such that $n = 2k$, and then $n^2 = 4k^2$, so n^2 is divisible; by 4, so the conclusion "'n^2 is divisible by 4" is true. So the sentence is true for every n.)
- $n(n + 1)(n + 2)$ is divisible by 6.
- If $n > 4$ then $n^2 > n+11$. (Reason: as we will see later, an implication "If A then B" is true if A is false or if B is true. Using this: if $n \leq 4$ then the implication "if $n > 4$ then $n^2 > n + 11$" is true because "'$n > 4$" is false. And if $n > 4$ then the implication "if $n > 4$ then $n^2 > n + 11$" is true because $n^2 > n + 11$' is true.)

On the other hand, you cannot say "$n^2 > 0$", because if you say that then the CAT will pick n to be 0, and you lose. □

Example 16. Suppose you say:

Let m, n be arbitrary natural numbers.

What can you say after that, being sure that it is true?

Certainly, you cannot say that $m = n$, because m and n could be different.

And you cannot say that $m \neq n$, because m and n could be equal.

And you cannot say that $m > n$, because m could be smaller than n.

But here are a few things you *can* say:

- $m + n \geq 2$. (Reason: $m \geq 1$ and $n \geq 1$, so $m + n \geq 2$.)
- $m.n$ is a natural number.
- $(m + n)^2 = m^2 + 2m + n^2$.
- $(m + n)^3 = m^3 + 3m^2n + 3mn^2 + n^3$.
- $m^2 - n^2 = (m - n)(m + n)$.
- $n + n^2$ and $m + m^2$ are even.
- Either $m > n$ or $m = n$ or $m < n$. □

6.1.7 Universal quantifiers and arbitrary things

Suppose you want to make sure (that is, prove) that something is true for ***all*** the members of some set S. For example, you may want to make sure that every student in a class knows that there is an exam next Tuesday.

You could do this in two ways:

1. You can use the ***exhaustive search method***: chack, one by one, all the memers of S, and verify that they all know about the exam.
2. You can use ***general reasoning***: you try to come up with an ***argument*** that shows that every student knows about the exam. (For example: maybe you have sent an e-mail to a mailing list of all the students, telling them about the exam. And yyou are sure that all the students get the messages to this mailing list, and that they all read them. Then you can be sure that they all know about the exam.)

If the set S is very large then it may be very difficult to use the exhaustive search method. And if the set is infinite then using exhaustive search is impossible. And this is the situation we encounter most of the time in

Mathematics: the sets S about we want to make sure that statements of the form "$P(x)$ is true for every member x of S" are usually infinite, or finite but very large. So the only way to prove that something is true for all members of some set S is to use ***reasoning***:

This is why, in order to prove universal sentences $(\forall x \in S)P(x)$, we use the following method:

- we imagine that we have an arbitrary member x of S,
- we reason about x, prove facts about x,
- and, maybe, eventually, we prove that $P(x)$, the fact about x that we wanted to make sure is true, is indeed true.

If we can do that for an ***arbitrary*** member of S, then we have established that $P(x)$ is true for every $x \in S$, that is, that $(\forall x \in S)P(x)$. ("$(\forall x \in S)P(x)$" is a "universally quantified sentence". We will study such sentences in great detail in Section 8, on page 93.)

The method for proving universally quantified sentences $(\forall x \in S)P(x)$ by proving that $P(x)$ is true for an arbitrary member x of S is the ***Rule for proving universal sentences***, that we will call $\boxed{\text{Rule } \forall_{prove}}$, This rule will be discussed in section 8.5, on page 105 below.

Problem 33. Indicate whether each of the following statements about n is true for an arbitrary integer n. If the answer is "yes", prove it. If the answer is "no", prove it by giving a counterxample, that is, a particular value of n for which the statement is false.

1. n is even.
2. n is even or n is odd.
3. n is even and n is odd.
4. n is even or $n + 1$ is even.
5. $n(n + 1)$ is even.
6. $n(n + 1)(n + 2)$ is divisible by 3.
7. $n(n + 1)(n + 2)$ is divisible by 6.
8. $n^2 > 0$.
9. $n^2 \geq 0$.

10. $n(n+1) \geq 0$.
11. $(\forall m \in \mathbb{Z})(n < m \implies n^2 < m^2)$.
12. $(\forall m \in \mathbb{Z})(n > m \implies n^2 > m^2)$.
13. $(\forall m \in \mathbb{Z})(n = m \implies n^2 = m^2)$.
14. $(\forall m \in \mathbb{Z})(n^2 = m^2 \implies n = m)$.

7 Dealing with equality

Throughout these notes, the symbols "=" and "≠" will be used.

- The symbol "=" is read as "is equal to".
- The symbol "≠" is read as "is not equal to".

The meaning of "=" in mathematics is quite simple: if a and b are any two things, then "$a = b$" (read as "a is equal to b", or "a equals b") means that a and b are the same thing.

Example 17.

- The sentence "$3 = 2 + 1$" is read as "three is equal to two plus one".
- The sentence "$3 = 2 + 2$" is read as "three is equal to two plus two".
- The sentence "$3 \neq 2 + 1$" is read as "three is not equal to two plus one".
- The sentence "$3 \neq 2 + 2$" is read as "three is not equal to two plus two".
- The sentences "$3 = 2 + 1$" and "$3 \neq 2 + 2$" are true.
- The sentences "$3 = 2 + 2$" and "$3 \neq 2 + 1$" are false. □

7.1 The substitution rule (Rule SEE, a.k.a. Rule $=_{use}$) and the axiom $(\forall x)x = x$

There are two basic facts you need to know about equality.

THE TWO BASIC FACTS ABOUT EQUALITY

First, there is the ***substitution rule***, which tells you that in a proof you can always "substitute equals for equals":

RULE SEE (substitution of equals for equals): If in a step of a proof you have an equality $s = t$ or $t = s$, and in another step you have a sentence P, then you can write as a step any statement obtained by substituting t for s in one or several of the occurrences of s in P.

The second thing you need to know is the following axiom:

EQUALITY AXIOM (*The "everything is equal to itself" axiom*):

$$x = x \text{ for every } x\,.$$

Example 18. In the sentence "$2 + 2 = 4$", the symbol "2" occurs twice. Suppose you have "$2 + 2 = 4$" as one of the steps in a proof. And suppose that in another step you have "$1 + 1 = 2$". Then you can substitute "$1 + 1$" for "2" in the first occurrence of "2" in the sentence "$2 + 2 = 4$", thus getting "$(1 + 1) + 2 = 4$". Or you can substitute "$1 + 1$" for "2" in the second occurrence of "2" in "$2 + 2 = 4$", thus getting "$2 + (1 + 1) = 4$". Or you can substitute "$1 + 1$" for "2" in both occurrences of "2" in "$2 + 2 = 4$", thus getting "$(1 + 1) + (1 + 1) = 4$". Or you can substitute "$1 + 1$" for "2" in none of occurrences, in which case you get back "$2 + 2 = 4$". □

Example 19. The following are true thanks to the equality axiom:

1. $3 = 3$,
2. $(345 + 1,031) \times 27 = (345 + 1,031) \times 27$,
3. Jupiter=Jupiter[34]
4. $\pi = \pi$.
5. My uncle Billy=My uncle Billy. □

[34]But you have to be ***very*** careful here! There are at least three different things named "Jupiter": a planet, a Roman god, and a Mozart symphony. When you write "Jupiter=Jupiter", you have to make sure that the two "Jupiter" in the equation have the same meaning. It would be false if you said that the planet Jupiter is the same as the Roman god Jupiter!

7.2 Equality is reflexive, symmetric, and transitive

Most textbooks will tell you that equality has the following three properties:

I. Equality is a ***reflexive*** relation. That is:

$$\text{for every } x\,,\ \ x = x\,. \tag{7.36}$$

II. Equality is a ***symmetric*** relation. That is:

$$\text{for every } x\,,\, y\,,\ \text{if } x = y \text{ then } y = x\,. \tag{7.37}$$

III. Equality is a ***transitive*** relation. That is:

$$\text{for every } x\,,\, y\,,\, z\,,\ \text{if } x = y \text{ and } y = z \text{ then } x = z \tag{7.38}$$

And, in addition, they will also tell you that the following important property holds:

IV. ***If two things are equal to a third thing then they are equal to each other.*** That is,

$$\text{for every } x\,,\, y\,,\, z\,,\ \text{if } x = z \text{ and } y = z \text{ then } x = y\,. \tag{7.39}$$

We could have put these properties as axioms, but we are not doing that because all these facts can easily be proved from our two basic facts about equality.

Theorem 12. *Facts I, II, III, and IV above follow from the two basic facts about equality described in the box on page 90 above.*

Proof. Fact I is exactly our Equality Axiom, so you don't need to prove it.

And now I am doing to do the proof of Fact II for you. So ***what you have to do is prove III and IV***.

Proof of Fact II.

Let x, y be arbitrary.

Assume $x = y$.

We want to prove that $y = x$.

By the Equality Axiom, $x = x$.

Since we have "$x = y$", Rule SEE tells us that, in the sentence "$x = x$", we can substitute "y" for any of the two occurrences of x in "$x = x$". So we choose to substitute "y" for the first of the two xs that occur in "$x = x$".

This yields $\boxed{y = x}$.

Since we have proved that $y = x$ assuming that $x = y$, we have shown that

$$\text{if } x = y \text{ then } y = x\,. \tag{7.40}$$

(This is because of Rule $\Longrightarrow_{prove}$, discussed later in these notes.)

Since we have proved (7.40) for arbitrary x, y, it follows that

$$\text{For all } x, y, \text{ if } x = y \text{ then } y = x\,. \tag{7.41}$$

(This is because of Rule $\forall_{prove}$, discussed later in these notes in section 8.5 on page 105.) This completes our proof. **Q.E.D.**

Proof of Facts III and IV. **YOU DO THEM.**

Problem 34. Write proofs of Fact III and Fact IV, following the model of the proof given here for Fact II. □

8 Universal sentences and how to prove and use them

A ***universal sentence*** is a sentence that says that something is true for every object x of a certain kind.

For example, the sentence

$$\text{every natural number is either even or odd} \tag{8.42}$$

says that every natural number has the property of being even or odd.

So this is a universal sentence.

Other examples of universal sentences are:

- Every natural number is an integer.
- Every real number has a square root[35].
- Every real number has a cube root[36].
- If n is any natural number then n is even or odd. □
- Every cow has four legs.
- Every cow has nine legs[37].
- All humans are thinking beings.
- All giraffes have a long neck.
- Every giraffe has a long neck.
- Every real number is positive[38].
- Every natural number can be written as the sum of three squares of integers[39].
- Every natural number can be written as the sum of four squares of integers[40].

[35] False!

[36] True!

[37] Sure, this one is false. But ***it is*** a universal sentence.

[38] This one is false.

[39] False again!

[40] This one, believe it or not, is true. But it is very hard to prove, and precisely for that reason, if you are interested in mathematics, I recommend that you read the proof. It is truly beautiful. The result is called "Lagrange's four squares theorem".

- Every integer is even[41].
- If a, b, c are integers, then if a divides b and c it follows that a divides $b + c$.

Universal sentences can always be rephrased is terms of "arbitrary things". For example, sentence (8.42) says

If n is an arbitrary natural number then n is either even or odd. (8.43)

We can say this in a more formal (and shorter) way by using the ***universal quantifier symbol***:

$$\forall$$

(This symbol is an inverted "A". The symbol is chosen to remind us that "$\forall$" stand for "for all".)

Precisely, the symbol is used as follows:

- Using the universal quantifier symbol, we form ***restricted universal quantifiers***, that is, expressions of the form

$$\big(\forall x \in S\big),$$

  where
  - x is a variable,
  - S is the name of a set.
- It is also possible to form ***unrestricted universal quantifiers***, that is, expressions of the form

$$\big(\forall x\big),$$

  where x is a variable,
- A restricted or unrestricted universal quantifier can be attached to a sentence by writing it before the sentence. This operation is called ***universal quantification***, and the result is a **universally quantified sentence**.

[41] Also false.

- So,

> If S is a set, and $P(x)$ is a statement involving the variable x, then
>
> $$(\forall x \in S)P(x)$$
>
> is a universally quantified sentence, obtained by universally quantifying the sentence $P(x)$.

> If $P(x)$ is a statement involving the variable x, then
>
> $$(\forall x)P(x)$$
>
> is a universally quantified sentence, obtained by universally quantifying the sentence $P(x)$.

8.1 How to read universal sentences

8.1.1 Sentences with restricted universal quantifiers

The universal sentence

$$(\forall x \in S)P(x)$$

can be read as follows:

- for every member x of S, $P(x)$ is true[42],

or as

- for every member x of S, $P(x)$,

or as

- for all members x of S, $P(x)$ is true,

or as

- for all members x of S, $P(x)$,

or as

- if x is an arbitrary member of S then $P(x)$ is true,

or as

- if x is an arbitrary member of S then $P(x)$.

[42] See Remark 7 below.

8.1.2 Sentences with restricted universal quantifiers

The universal sentence

$$(\forall x)P(x)$$

can be read as follows:

- for every x, $P(x)$ is true[43],

or as

- for every x, $P(x)$,

or as

- for all x, $P(x)$ is true,

or as

- for all x, $P(x)$,

or as

- if x is arbitrary then $P(x)$ is true,

or as

- if x is arbitrary then $P(x)$.

8.1.3 A recommendation

Of all these ways of reading "$(\forall x \in S)P(x)$" and "$(\forall x)P(x)$", ***I strongly recommend the ones involving "arbitrary"*** x, because once you get used to reading universal statements that way it becomes very clear how to go about proving them.

Remark 7. If A is any sentence, then saying "A is true" is just another way of asserting A. For example, saying that

$$\text{"all animals are made of cells}'' \text{ is true} \tag{8.44}$$

is just another way of saying

$$\text{all animals are made of cells}\,. \tag{8.45}$$

[43] See Remark 7 below.

Similarly, saying

$$P(n) \text{ is true} \tag{8.46}$$

is just another way of saying

$$P(n)\,. \tag{8.47}$$

This is why the sentence "$(\forall n \in \mathbb{Z})P(n)$" can be read either as "if n is an arbitrary integer then $P(n)$ is true", or as "if n is an arbitrary integer then $P(n)$". □

8.2 Using the universal quantifier symbol to write universal statements

8.2.1 What is formal language?

As we explained before, ***formal language*** is a language in which you use only formulas, and no words.

For example, you know from your early childhood how to take the English sentence "two plus two equals four" and say the same thing in formal language. i.e., with a formula. You just write

$$2 + 2 = 4\,. \tag{8.48}$$

We can say more complicated things in formal language by introducing more symbols. For example, here is the definition if "divisible" that we saw earlier:

DEFINITION Let a, b be integers. We say that a is divisible by b (or that b is a factor of a) if there exists an integer k such that $a = bk$. □

Then, we can agree to introduce the new symbol "|" to stand for "is a factor of", and write

$$b|a \tag{8.49}$$

instead of "b is a factor or a", or "a is divisible by b".

In particular, we can now say "x is even" in formal language, as follows: "$2|x$". So, for example the assertion that "the sum of two even integers is even" becomes, in formal language:

$$(\forall a \in \mathbb{Z})(\forall b \in \mathbb{Z})\Big((2|a \wedge 2|b) \Longrightarrow 2|a+b\Big)\,. \tag{8.50}$$

Can you say more complicated things in formal language? For example, can you rewrite the English sentence

(#)	If we take any two real numbers and compute the square of their sum, then you get the same result as when you add the squares of the two numbers plus twice their product.

in formal language?

You know since high school that you can take a big part of (#) and rewrite it in formal language. The trick is to ***give names*** to the two integers that you want to talk about. Then you can write

(#1)	If we take any two real numbers and call them a and b, then $$(a+b)^2 = a^2 + b^2 + 2ab\,,$$

or

(#2)	If a, b are arbitrary real numbers, then $$(a+b)^2 = a^2 + b^2 + 2ab\,.$$

Naturally, you could use any names you want, For example, you could equally well have written

(#3)	If x, y are arbitrary real numbers, then $$(x+y)^2 = x^2 + y^2 + 2xy\,.$$

or

(#4)	If we take any two real numbers and call them x and y, then $$(x+y)^2 = x^2 + y^2 + 2xy\,.$$

Sentences (#1), (#2), (#3), (#4) are statements in ***semiformal language***: they are a mixture of formal language and ordinary English.

These statements are universal sentences. And now you have learned how to ***formalize***[44] universal statements. So you can write

(#5)	$(\forall a \in \mathbb{R})(\forall b \in \mathbb{R})(a+b)^2 = a^2 + b^2 + 2ab\,.$

or

[44] that is, how to say in formal language

(#6)	$(\forall x \in \mathbb{R})(\forall y \in \mathbb{R})(x+y)^2 = x^2 + y^2 + 2xy\,.$

Statements (#5) and (#6) are ***formal sentences***, that is, formulas with no words.

8.2.2 The road to full formalization.

What we have done is get started moving towards full formalization.

You started doing this in your childhood, when you learned how to formalize "two plus two equals four" by writing "$2 + 2 = 4$".

And now you have learned how to formalize more complicated sentences, Using the universal quantifier symbol, you are now able to say many more things in formal language.

Example 20. Suppose you wanted to say "every natural number is positive". You can write

$$(\forall n \in \mathbb{N})n > 0\,. \tag{8.51}$$

This is a formula, that is, a sentence in formal language. □

Example 21. Although we do not know yet how to write something like

(#7)	If we have any two integers, when say that the first one is divisible by the second one what we mean is that there exists an integer that multiplied by the second one results in the first one.

in full formal language, we are able, using what we know so far, to go a long way, and rewrite (#7) in semiformal language, with very few words, i.e., getting very close to a fully formal sentence. We can write

(#8)	$(\forall a \in \mathbb{Z})(\forall b \in \mathbb{Z})$("$a\|b$" means "there exists k such that $k \in \mathbb{Z}$ and $b = ak$.")

□

Example 22. Let us say "If a, b, c are integers, then if a divides b and c it follows that a divides $b + c$" in semiformal language.

We can say:

$$(\forall a \in \mathbb{Z})(\forall b \in \mathbb{Z})(\forall c \in \mathbb{Z})\Big(\text{if } a|b \text{ and } a|c \text{ then } a|b+c\Big)\,, \tag{8.52}$$

which is, again, a sentence in semiformal language. □

Later, when we learn how to say "means", "there exists", "if ... then" and "and", we will be able to say (#8) and (8.52) in fully formal language, as follows:

- We can translate (#8) into fully formal language as

$$(\forall a \in \mathbb{Z})(\forall b \in \mathbb{Z})(a|b \iff (\exists k \in \mathbb{Z})b = ak)\,. \tag{8.53}$$

- We can translate (8.52) into fully formal language as

$$(\forall a \in \mathbb{Z})(\forall b \in \mathbb{Z})(\forall c \in \mathbb{Z})\Big((a|b \wedge a|c) \implies a|b+c\Big)\,, \tag{8.54}$$

8.3 Open and closed variables and quantified sentences

Let us recall that

> A free variable is a letter (or string of symbols) that is "unattached", in the sense that it has no particular value, and is free to be assigned any value we want.
> A bound variable is a variable that has been assigned a specific value, by means of a ***value declaration***.
> We can turn a free variable into a temporary constant by ***declaring its value.***

Let me add a couple of points to that:

- Free variables are also called open variables.
- Bound variables are also called closed variables.

  (They are called "bound" variables because they are "bound", attached to a value, by contrast with free variables, that are free to be assigned any value because they do not have a value already assigned to them. And they are called "closed" because they are not open to be assigned a value, since they already have one.)

- ***A value declaration is valid until it expires***. When the value declaration expires, the variable becomes free again, and you can assign a new value to it.

Example 23. Here is an example of declaring a value for a variable, and of making that declaration expire. You could write:

1. Let $x = \frac{1+\sqrt{5}}{2}$.
2. Then $x^2 = 1 + x$.
3. Now suppose, instead, that $x = \frac{1-\sqrt{5}}{2}$.
4. Then it is also true that $x^2 = 1 + x$.

Here, step 1 assigns the value $\frac{1+\sqrt{5}}{2}$ to the variable, so this variable, which until then was open, is now attached to the value $\frac{1+\sqrt{5}}{2}$, so x is bound, no longer free.

But then, in step 3, we are ssigning a new value to x, which means that the previous value declaration has expired. The fact that the previous value declaration has expired is signaled by the word "now"', and reinforced by the word "instead".

Notice that if you had written

1. Let $x = \frac{1+\sqrt{5}}{2}$.
2. Then $x^2 = 1 + x$.
3. Let $x = \frac{1-\sqrt{5}}{2}$.
4. Then it is also true that $x^2 = 1 + x$.

this would have been confusing for many readers, because they would have wondered: "wasn't x equal to $\frac{1+\sqrt{5}}{2}$? How come suddenly it seems to have a different value?"

The words "now" and "instead" make it crystal clear to the reader that the first value declaration has just expired and we are free to assign to x a new value if we so desire. □

8.4 A general principle: two rules for each symbol

Every time we introduce a new symbol, we need two rules telling us how to work with it:

- We need a rule that tells us how to ***use*** statements involving that symbol.

and

- We need a rule that tells us how to ***prove*** statements involving that symbol.

Example 24. Let us look at the new symbol "|" ("divides") that we introduced in Part I of these notes. What is the "use" rule'? What is the "prove" rule?

The "use" rule is:

> If you get to a point in a proof where you have a statement
>
> $$a|b\,,$$
>
> then you can go from this to
>
> We may pick an integer k such that $b = ak$.

And the "prove" rule is:

> If you get to a point in a proof where you have integers a, b, c and you know that
>
> $$b = ak\,,$$
>
> then you can go from this to
>
> $$a|b\,.$$

These rules are just another way of stating the definition of "divides". □

8.4.1 Naming sentences

Sentences are also things that we can talk about, so we can give them names.

One common way mathematicians use to name sentences is to give a sentence a capital letter name, such as A, or B, or P, or Q, or S.

So we could talk about the sentence "x eats grass" by giving it a name, that is, by picking a capital letter and declaring its value to be this sentence.

We could do this by writing

Let P be the sentence "x eats grass".

However, there is a much more convenient way to do this: ***If a sentence has an open variable, we include this open variable in the name of the sentence, thus signaling to the reader that the sentence contains that open variable.***

So, for example, a good name for the sentence "x eats grass" could be $P(x)$ (or $A(x)$, or $S(x)$, etc.). We could declare the value of the variable $P(x)$ by saying

(*) Let $P(x)$ be the sentence "x eats grass".

An important convention about names of sentences is this: suppose we want to talk about the sentence obtained from $P(x)$ by substituting (i.e., "plugging in") the name of a particular thing for the open variable x. If we already have a name for that thing, say "a", then the name of the sentence arising from the substitution is $P(a)$.

So, for example, after we make the value declaration (*), then "P(Suzy)" is the name of the sentence "Suzy eats grass".

What if you have a sentence with, say, two or more open variables? You do the same thing: if, for example, you want to give a name to the sentence "x told y that z does not like w", you can call that sentence $P(x, y, z, w)$. You could make the value declaration

> Let $P(x, y, z, w)$ be the sentence "x told y that z does not like w"."

And then,

- If you want want to talk about the sentence "Alice told Jim that Bill does not like Mary", then that sentence would have the name $P(\text{Alice}, \text{Jim}, \text{Bill}, \text{Mary})$.
- If you want want to talk about the sentence "Alice told Jim that Bill does not like her" (that is, does not like Alice), that sentence would be called $P(\text{Alice}, \text{Jim}, \text{Bill}, \text{Alice})$.
- If you want want to talk about the sentence "Alice told Jim that Bill does not like him" (that is, does not like Jim), that sentence would be called $P(\text{Alice}, \text{Jim}, \text{Bill}, \text{Jim})$.
- And, if, for some reason, you want to talk about the sentence with two open variables "x told y that Bill does not like Mary", that sentence would be $P(x, y, \text{Jim}, \text{Mary})$.

8.4.2 Universal sentences bound variables but at the end let them free

If $P(x)$ is a sentence with the open variable x, and C is a set, then the sentence

$$(\forall x \in C)P(x)$$

should be read as

> Let x be an arbitrary member of C; then $P(x)$ is true; and now the value declaration of "x" expires, and x is a free variable again.

Why do we want to do this?

The reason is that the value declaration ("Let x be an arbitrary member of C") was made for the sole purpose of explaining which condition this arbitrary member of C is supposed to satisfy. Once this has been explained, there is no need to keep the variable x bound forever. It is better to let it be free again, so that the next time we need a variable for something, we can use x.

So, for example, when I explain to you that

(F) If x is an arbitrary integer then $(x+1)^2 = x^2 + 2x + 1$,

the important thing that I want you to remember is that "if you take an integer, add one to it, and square the result, then what you get is the sum of the square of your integer, plus two times it, plus one". There is no need for you to remember, in addition, the name that I used for that integer for the purpose of explaining Fact (F) to you. You should not have to waste any time or effort trying to remember "was that fact that was explained to me about x? Or was it about y? Or was it about n?" There is not need for you to remember that, because ***it does not matter which variable was used***. And, more importantly: ***Fact (F) is not really about a specific integer called x. It is a fact about an arbitrary integer, and it does not matter whether you call it x, or y, or z, or n, or α, or β, or $\diamond$, or even "Suzy" or "my uncle Jimmy". The letter x is used as a device within the conversation in which you explain Fact (F) to me, and once this conversation is over we can forget about x.***

Example 25. Suppose you have written, in a proof:

$$(\forall n \in \mathbb{Z}) n(n+1) \text{ is even}. \tag{8.55}$$

Can you write, in the next step of your proof:

Since $n(n+1) = n + n^2$, it follows that $n + n^2$ is even. ?

The answer is ***no***. Why? Because after the end of the sentence (8.55), n is a free variable again, so it does not have a value, so when you use "n" in the next step, nobody knows what you are talking about, so what you wrote is meaningless, so it's not acceptable.

Suppose you want to go from (8.55) to

$$(\forall n \in \mathbb{Z}) n + n^2 \text{ is even}. \tag{8.56}$$

How can you do that? The answer is: you use the rules for using and proving universal sentences. But ***you do it correctly***. And for that you need to read the next section. □

8.5 Proving and using universal sentences (Rules $\forall_{prove}$ and $\forall_{use}$)

Now that we know that for every new symbol we introduce we need a "use" rule and a "prove" rule, it is natural to ask: *What are the "use" rule and the "prove" rule for the universal quantifier symbol $\forall$?"*

Both are very simple, very natural rules.

Here is the "use" rule:

The rule for using universal sentences
(Rule $\forall_{use}$, also known as
the "universal specialization rule")

- If you have proved
$$(\forall x)P(x)\,,$$
and you have an object called a, then you can go to $P(a)$.
- If you have proved
$$(\forall x \in S)P(x)\,,$$
and you have an object called a for which you know that $a \in S$, then you can go to $P(a)$.

The reason Rule $\forall_{use}$ is called called the ***universal specialization rule***, is that the rule says that if a statement is true in general (that is, for all things that belong to some set S), then it is true in each special case (that is, for a particular thing that belongs to S).

Example 26. If you know that $(\forall x)x = x$, then you can conclude from that, using Rule $\forall_{use}$, that
$$3 = 3\,,$$
and that
$$5 + 3 = 5 + 3\,.$$

Example 27. Suppose you know that

(&) All cows eat grass.

and that

(&&) Suzy is a cow.

Then, from (&) and (&&) you can conclude, thanks to the specialization rule, that

(&&) Suzy eats grass.

In formal language. you would say this as follows: Let $P(x)$ be the sentence "x eats grass", and let C be the set of all cows. Then $P(\text{Suzy})$ is the sentence "Suzy eats grass". And (&) says

(&') $(\forall x \in C)P(x)\,,$

whereas (&&) says

(&&') $\text{Suzy} \in C.$

So we are precisely in the situation where we can apply the rule for using universal sentences, and conclude that $P(\text{Suzy})$, that is that Suzy eats grass. □.

And here is the "prove" rule:

The rule for proving universal sentences

- To prove $(\forall x)P(x)$, you start by writing

  Let x be arbitrary,

  and then prove $P(x)$

  If you manage to do that, then you are allowed to write

  $$(\forall x)P(x)$$

  in the next step of your proof.

- To prove $(\forall x \in S)P(x)$, you start by writing

  Let x be an arbitrary member of S,

  and then prove $P(x)$

  If you manage to do that, then you are allowed to write

  $$(\forall x \in S)P(x)$$

  in the next step of your proof.

This rule is also called the ***generalization rule***, because it says that if you can prove a statement for an arbitrary object that belongs to a set S then you can "generalize", i.e., conclude that the statement is true in general, for all members of S.

8.6 An example: Proof of the inequality $x + \frac{1}{x} \geq 2$

Let us illustrate the use of the proof rules for universal quantifiers with an example. We will first present a version of the proof with lots of comments. The comments are explanations to help the reader follow what is going on, but are not really necessary for the proof. We will then present another, much shorter version, in which the comments are omitted.

Theorem 13. *If x is a positive*[45] *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, WITH LOTS OF COMMENTS. (The comments are in Italics.)
The assertion we want to prove is a universal sentence, so we are going to use Rule $\forall_{prove}$. For that purpose, we imagine we have in our hands an arbitrary real number called x, and we work with that number.

> Let x be an arbitrary real number.
>
> *Now we want to prove that $x > 0 \Longrightarrow x + \frac{1}{x} \geq 2$. This is an implication. So we are going to apply Rule $\Longrightarrow_{prove}$. For that purpose, we assume that the premise of our implication is true, i.e., that $x > 0$. The reason for this is as follows: x is an arbitrary real number, so x could be any ral number, and in particular x could be positive, negative, or zero. If x is not positive, then the implication is true, because an implication whose premise is false is true. So all we need is to look at the cases when $x > 0$, and prove in that case that $x + \frac{1}{x} \geq 2$.*
>
> > Assume that $x > 0$.
> >
> > We want to prove that
> >
> > $$x + \frac{1}{x} \geq 2\,. \tag{8.57}$$

[45] 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".

We will prove this by contradiction.

Assume that (8.57) is not true.

Then

$$x + \frac{1}{x} < 2\,. \tag{8.58}$$

We now use a fact from real number arithmetic, namely, that if we multiply both sides of a true inequality by a positive real number then the result is a true inequality, that is:

$$(\forall a \in \mathbb{R})(\forall b \in \mathbb{R})(\forall c \in \mathbb{R})\Big((a < b \wedge c > 0) \Longrightarrow ac < bc\Big)\,. \tag{8.59}$$

In our case. we can use Rule $\forall_{use}$ to plug in $x + \frac{1}{x}$ for a, 2 for b, and x for c in (8.59), and get

$$\Big(x + \frac{1}{x} < 2 \wedge x > 0\Big) \Longrightarrow \Big(x + \frac{1}{x}\Big)x < 2x\,. \tag{8.60}$$

Since $x + \frac{1}{x} < 2 \wedge x > 0$ is true (because we are assuming that $x + \frac{1}{x} < 2$ and that $x > 0$), we can apply Rule $\Longrightarrow_{use}$ to conclude that $\Big(x + \frac{1}{x}\Big)x < 2x$. But $\Big(x + \frac{1}{x}\Big)x = x^2 + 1$, so we have shown that $x^2 + 1 < 2x$. Summarizing:

Since $x > 0$, we can multiply both sides of (8.58) by x, getting

$$x^2 + 1 < 2x\,. \tag{8.61}$$

Now we use another fact from real number arithmetic, namely, that if we add a real number to both sides of a true inequality, then the result is a true inequality, that is:

$$(\forall a \in \mathbb{R})(\forall b \in \mathbb{R})(\forall c \in \mathbb{R})(a < b \Longrightarrow a + c < b + c)\,. \tag{8.62}$$

In our case. we can use Rule $\forall_{use}$ to plug in $x^2 + 1$ for a, $2x$ for b, and $-2x$ for c in (8.62), and get

$$x^2 + 1 - 2x < 2x - 2x,\,. \tag{8.63}$$

Since $2x - 2x = 0$, we can conclude that $x^2 + 1 - 2x < 0$. Summarizing:

We add $-2x$ to both sides, and get

$$x^2 + 1 - 2x < 0\,. \tag{8.64}$$

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

$$(x-1)^2 < 0\,. \tag{8.65}$$

Now we use a third fact from real number arithmetic, namely, that the square of every real number is nonnegative, that is:

$$(\forall u \in \mathbb{R})u^2 \geq 0\,. \tag{8.66}$$

We use Rule $\forall_{use}$ to plug in $x-1$ for u, and get

$$(x-1)^2 \geq 0\,. \tag{8.67}$$

Next, we use a fourth fact from real number arithmetic, namely, that if a real number is nonnegative then it it is not negative[46], that is:

$$(\forall u \in \mathbb{R})(u \geq 0 \Longrightarrow \sim u < 0)\,. \tag{8.68}$$

It then follows from (8.67) that

$$\sim (x-1)^2 < 0\,. \tag{8.69}$$

From (8.65) and (8.69), we get

$$(x-1)^2 < 0 \wedge \Big(\sim (x-1)^2 < 0 \Big)\,. \tag{8.70}$$

So we have proved a contradiction.

We have proved that a world in which the inequality $x + \frac{1}{x} > 2$ is not true is an impossible world. Hence
$x + \frac{1}{x} > 2$.

We have proved that $x + \frac{1}{x} > 2$ assuming that $x > 0$. Hence Rule $\Longrightarrow_{prove}$ allows us to conclude that

$$x > 0 \Longrightarrow x + \frac{1}{x} \geq 2\,. \tag{8.71}$$

Finally, we have proved (8.71) for an arbitrary real number x. Hence

$$(\forall x \in \mathbb{R})(x > 0 \Longrightarrow x + \frac{1}{x} \geq 2)\,. \tag{8.72}$$

Q.E.D.

[46]Remember that: "positive" means "> 0", "negative" means "< 0", "nonnegative" means "≥ 0", and "nonpositive" means "≤ 0".

THE SAME PROOF, WITHOUT THE COMMENTS.

Let x be an arbitrary real number.

Assume that $x > 0$.
We want to prove that

$$x + \frac{1}{x} \geq 2\,. \tag{8.73}$$

Assume that (8.73) is not true.
Then

$$x + \frac{1}{x} < 2\,. \tag{8.74}$$

Since $x > 0$, we can multiply both sides of (8.74) by x, getting

$$x^2 + 1 < 2x\,. \tag{8.75}$$

We add $-2x$ to both sides, and get

$$x^2 + 1 - 2x < 0\,. \tag{8.76}$$

But $x^2 + 1 - 2x = (x-1)^2$. So

$$(x-1)^2 < 0\,. \tag{8.77}$$

Now we use the fact that the square of every real number is nonnegative, that is:

$$(\forall u \in \mathbb{R}) u^2 \geq 0\,. \tag{8.78}$$

We use Rule $\forall_{use}$ to plug in $x - 1$ for u, and get

$$(x-1)^2 \geq 0\,. \tag{8.79}$$

Then

$$\sim (x-1)^2 < 0\,. \tag{8.80}$$

From (8.77) and (8.80), we get

$$(x-1)^2 < 0 \wedge \Big(\sim (x-1)^2 < 0 \Big) \,. \tag{8.81}$$

So we have proved a contradiction.

Hence $x + \frac{1}{x} > 2$.

We have proved that $x + \frac{1}{x} > 2$ assuming that $x > 0$. Hence Rule $\Longrightarrow_{prove}$ allows us to conclude that

$$x > 0 \Longrightarrow x + \frac{1}{x} \geq 2 \,. \tag{8.82}$$

Finally, we have proved (8.80) for an arbitrary real number x. Hence

$$(\forall x \in \mathbb{R})(x > 0 \Longrightarrow x + \frac{1}{x} \geq 2) \,. \tag{8.83}$$

Q.**E**.**D**.

THE SAME PROOF, IN A MUCH SHORTER VERSION.

Let x be an arbitrary real number.

Assume that $x > 0$. We want to prove that

$$x + \frac{1}{x} \geq 2\,. \tag{8.84}$$

Assume that (8.84) is not true. Then

$$x + \frac{1}{x} < 2\,. \tag{8.85}$$

Since $x > 0$, (8.85) impliues

$$x^2 + 1 < 2x\,. \tag{8.86}$$

Therefore

$$x^2 + 1 - 2x < 0\,. \tag{8.87}$$

But $x^2 + 1 - 2x = (x-1)^2$. So

$$(x-1)^2 < 0\,. \tag{8.88}$$

On the other hand.

$$(x-1)^2 \geq 0\,. \tag{8.89}$$

Clearly, (8.88) and (8.89) lead to a contradiction.

Hence
$x + \frac{1}{x} > 2$.

Therefore

$$(\forall x \in \mathbb{R})(x > 0 \Longrightarrow x + \frac{1}{x} \geq 2)\,. \tag{8.90}$$

Q.E.D.

8.6.1 A few more examples of proofs involving universal sentences

Theorem 14. *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 35. Prove Theorem 14.

Problem 36. Explain what is wrong with the following proof of Theorem 14.

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.**

Theorem 15. *If x, α, β are positive real numbers then*

$$\alpha x + \frac{\beta}{x} \geq 2\sqrt{\alpha\beta}\,.$$

(In formal language: $(\forall \alpha \in \mathbb{R})(\forall \beta \in \mathbb{R})(\forall x \in \mathbb{R})\Big((\alpha > 0 \wedge \beta > 0 \wedge x > 0) \Longrightarrow \alpha x + \frac{\beta}{x} \geq 2\sqrt{\alpha\beta}\Big)$*.)*

I am going to give you two proofs. The first one follows the same pattern as the proof of Theorem 13. The second one, much shorter, uses Theorem 13.

FIRST PROOF.

Let α, β, x be arbitrary positive real numbers[47].

[47] In this one step I am conflating six real steps: let α be an arbitrary real number, let β be an arbitrary real number, let x be an arbitrary real number, assume $\alpha > 0$, assume $\beta > 0$, assume $x > 0$.

Let $q = 2\sqrt{\alpha\beta}$, so $\frac{q^2}{4\alpha} = \beta$..

Assume $\sim \alpha x + \frac{\beta}{x} \geq q$.

Then $\alpha x + \frac{\beta}{x} < q$.

Therefore $\alpha x^2 + \beta < qx$.

Hence $\alpha x^2 - qx + \beta < 0$.

But

$$\begin{aligned} \alpha x^2 - qx + \beta &= \alpha x^2 - 2\sqrt{\alpha} x \frac{q}{2\sqrt{\alpha}} + \beta \\ &= \alpha x^2 - 2\sqrt{\alpha} x \frac{q}{2\sqrt{\alpha}} + \frac{q^2}{4\alpha} - \frac{q^2}{4\alpha} + \beta \\ &= \left(\sqrt{\alpha} x - \frac{q}{2\sqrt{\alpha}}\right)^2 \\ &\geq 0\,. \end{aligned}$$

So we obtain a contradiction, and then we can conclude that $\alpha x + \frac{\beta}{x} \geq q$, i.e. that

$$\alpha x + \frac{\beta}{x} \geq 2\sqrt{\alpha\beta}\,.$$

Q.E.D.

SECOND PROOF. Let us try to write $\alpha x + \frac{\beta}{x}$ as $p\Big(u + \frac{1}{u}\Big)$ for some positive u, and use the fact that $u + \frac{1}{u} \geq 2$. Let $x = hu$, where h and u are to be determined later.

Then $\alpha x + \frac{\beta}{x} = \alpha h u + \frac{\beta}{hu}$. If we could make $\alpha h = \frac{\beta}{h}$, we would get

$$\begin{aligned} \alpha x + \frac{\beta}{x} &= \alpha h u + \frac{\beta}{hu} \\ &= \alpha h u + \alpha h \frac{1}{u} \\ &= \alpha h\Big(u + \frac{1}{u}\Big)\,, \end{aligned}$$

as desired.

So we need to chose h such that $\alpha h = \frac{\beta}{h}$, that is, such that $h = \sqrt{\frac{\beta}{\alpha}}$.

With this choice of h, we get

$$\begin{aligned} \alpha x + \frac{\beta}{x} &= \alpha h\Big(u + \frac{1}{u}\Big) \\ &\geq 2\alpha h \\ &= 2\alpha\sqrt{\frac{\beta}{\alpha}} \\ &= 2\sqrt{\alpha\beta}\,. \end{aligned}$$

Q.E.D.

8.6.2 * The inequality $\frac{x^n}{n} - ax \geq -\frac{n-1}{n}a^{\frac{n}{n-1}}$: a proof using Calculus

Theorem 16. *Let a and b be positive real numbers, and let n be a positive integer. Then*

$$ab \leq \frac{1}{n}\Big(a^n + (n-1)b^{\frac{n}{n-1}}\Big)\,. \tag{8.91}$$

Remark 8. For $n = 2$, inequality (8.91) says that

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

which is Theorem 14.

So (8.91) is a generalization of Theorem 14. □

Proof of Theorem 16. We will use Calculus.

Let a, b be arbitrary positive real numbers.

Define a function f by letting

$$f(x) = \frac{x^n}{n} - bx \text{ for } x \in \mathbb{R}\,,\ x \geq 0\,.$$

We would like to find the value of x where f has its minimum value of f for all positive x. That is, we would like to find a positive real number c such that $f(c) \leq f(x)$ for all positive x.

For this purpose, we compute the derivative f' of f.

We have

$$f'(x) = x^{n-1} - b \text{ for every } x \in \mathbb{R}\,.$$

Let $c = b^{\frac{1}{n-1}}$. Then $c^{n-1} = b$, so $f'(c) = c^{n-1} - b = 0$.

This means that c is a candidate for our minimum. That is, it is possible that c is where f has its minimum value, in which case it would follow that

$$f(x) \geq f(c) \text{ for all } x \in \mathbb{R} \text{ such that } x > 0\,. \tag{8.92}$$

We now prove (8.92) rigorously

If $0 < x < c$, then $x^{n-1} < c^{n-1} = b$, so $x^{n-1} - b < 0$, so $f'(x) < 0$.

This means that the function f is decreasing for $0 < x < c$. So $f(x) \geq f(c)$ for $0 < x < c$.

If $x > c$, then $x^{n-1} > c^{n-1} = b$, so $x^{n-1} - b > 0$, so $f'(x) > 0$.

This means that the function f is increasing for $x > c$. So $f(x) \geq f(c)$ for $x > c$.

We have shown that $f(x) \geq f(c)$ when $0 < x < c$ and when $x > c$. And clearly $f(x) = f(c)$ when $x = c$. Hence (8.92) is true.

It follows from (8.92) that for every positive $x \in \mathbb{R}$ we have $f(x) \geq f(c)$, that is,

$$\frac{x^n}{n} - bx \geq \frac{c^n}{n} - bc\,. \tag{8.93}$$

Since (8.93) holds for every positive x, we can use it for $x = a$, thereby obtaining

$$\frac{a^n}{n} - ab \geq \frac{c^n}{n} - bc\,. \tag{8.94}$$

Since $c = b^{\frac{1}{n-1}}$ and $c^{n-1} = b$, we have

$$\begin{aligned}
\frac{c^n}{n} - bc &= \frac{b^{\frac{n}{n-1}}}{n} - b \times b^{\frac{1}{n-1}} \\
&= \frac{b^{\frac{n}{n-1}}}{n} - b^{1+\frac{1}{n-1}} \\
&= \frac{b^{\frac{n}{n-1}}}{n} - b^{\frac{n}{n-1}} \\
&= \Big(\frac{1}{n} - 1\Big) b^{\frac{n}{n-1}} \\
&= -\frac{n-1}{n} b^{\frac{n}{n-1}}\,.
\end{aligned}$$

In view of (8.94), we get

$$\frac{a^n}{n} - ab \geq -\frac{n-1}{n} b^{\frac{n}{n-1}}\,, \tag{8.95}$$

that is,

$$\frac{a^n}{n} - ab + \frac{n-1}{n} b^{\frac{n}{n-1}} \geq 0\,, \tag{8.96}$$

from which it follows that

$$ab \leq \frac{a^n}{n} + \frac{n-1}{n} b^{\frac{n}{n-1}}\,, \tag{8.97}$$

that is,

$$ab \leq \frac{1}{n}\Big(a^n + (n-1)b^{\frac{n}{n-1}}\Big)\,, \tag{8.98}$$

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

9 Existential sentences

9.1 Existential quantifiers

- The symbol

$$\exists$$

is the ***existential quantifier symbol***.

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

"$(\exists x)$" is an ***unrestricted existential quantifier***,

and

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

- Existential 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"

Example 28. The sentence

$$(\exists x \in \mathbb{R}) x^2 = 2 \tag{9.99}$$

could be read as

> There exists an x belonging to the set of real numbers such that $x^2 = 2$.

But this is horrible! A much better way to read it is:

> There exists a real number x such that $x^2 = 2$.

An even better way is

> There exists a real number whose square is 2.

And the nicest way of all is

> 2 has a square root.

And you can also read (9.99) as:

> It is possible to pick a real number x such that $x^2 = 2$.

I strongly recommend this reading, because when you read an existential sentence this way it becomes clear that the next thing to do is to actually pick an x, that is, to apply the rule for using an existential sentence, i.e. Rule $\exists_{use}$ □

9.1.1 How not to read existential quantifiers

Students sometimes read an existential sentence such as

$$(\exists x \in \mathbb{R}) x^2 = 2) \tag{9.100}$$

as follows: *there exists a real number x and $x^2 = 2$.*

This is completely wrong, and should be avoided at all costs, because if you read an existential sentence that way you are going to be led to making lots of other mistakes.

Why is this wrong?

- If you read (9.100) as "there exists a real number x and $x^2 = 2$", then you give the impression that (9.100) makes two assertions:
  1. that there exists a real number,

2. that $x^2 = 2$.

- But (9.100) does not say that at all! What it does is make ***one*** assertion, namely, that there exists a real number x such that $x^2 = 2$. ("Such that" means "for which it is true that".)

If you are asked to prove (9.100) and you read is as "there exists a real number x and $x^2 = 2$", then you will think that you have to prove two things, namely, (1) that there exists a real number, and (2) that $x^2 = 2$. But what you have to prove is one thing: that it is possible to pick a real number whose square is 2.

The word "and" in this bad reading is particularly pernicious, because it makes you see two sentences where there is only one sentence. ***The quantifier*** $(\exists x \in \mathbb{R})$ ***is not a sentence.***

You can see this even more clearly if you read (9.100) as "for some real numbers x, $x^2 = 2$". It is clear that "for some real numbers x" is not a sentence. And it's nonsense to say "for some real numbers x and $x^2 = 2$".

Since "for some real numbers x" is another way to read the quantifier $(\exists x \in \mathbb{R})$, it should be clear that there is no "and" in such a quantifier,

9.1.2 Witnesses

A witness for an existential sentence $(\exists x)P(x)$ is an object a such that $P(a)$ is true.

A witness for an existential sentence $(\exists x \in S)P(x)$, is an object a such that $a \in S$ and $P(a)$ is true.

9.2 How do we work with existential sentences in proofs?

As you may have guessed, I am going to give you two rules, one for *proving* existential sentences, and one for *using* them. And the names of these rules are going to be—yes, you guessed it!—Rule $\exists_{prove}$ and Rule $\exists_{use}$.

9.2.1 The rule for using existential sentences (Rule $\exists_{use}$)

Rule $\exists_{use}$ says something very simple and natural: ***if you know that an object of a certain kind exists, then you can pick one and give it a name.***

In other words, ***if you know that*** $(\exists x)P(x)$ ***or that*** $(\exists x \in S)P(x)$***, then you are allowed to pick a witness and give it a name.***

Example 29. Suppose "$P(x)$" stands for "x eats grass", and C is the set of all cows. Suppose you know that

$$(\exists x \in C)P(x)\,, \tag{9.101}$$

that is, you know that there are grass-eating cows.

Then the thing you can do, according to Rule $\exists_{use}$, is pick a cow and give her a name.

So, for example, you could write

Pick a cow that eats grass and call her Suzy.

Or you could write

Let Suzy be a witness for the sentence (9.101,
so Suzy is a grass-eating cow.

or

Let Suzy be a grass-eating cow.

Example 30. Suppose you have a real number x and you know that

$$(\exists y \in \mathbb{R})y^5 - y^3 = x\,. \tag{9.102}$$

Then you can say, in the next step of your proof: :

Pick a witness for (9.102) and call it r, so $r \in \mathbb{R}$ and $r^5 - r^3 = 5$.

or you could write

Let r be a real number such that $r^5 - r^3 = 5$.

And you could even say

Let y be a real number such that $y^5 - y^3 = 5$.

□

Remark 9. When you pick a witness, as in the previous example, you can give it any name you want: you can call it r, k, m, u, $\hat{r}$, a, α, $\diamond$, ♣, Alice, Donald Duck, whatever.

You can even call it y, if you wish.

The key point is: ***the name you use cannot be already in use as the name of something else***.

So "y" qualifies as an acceptable name because, within the sentence "$(\exists y \in \mathbb{R})y^5 - y^3 = x$", y is a bound variable, but as soon as the sentence ends, "y" becomes a free variable, with no declared value, so you are allowed to use it.

However, I recommend that you do not use the same letter that appeared in the existential quantifier. □

There is, however, one thing that is absolutely forbidden:

> ***You cannot give the new object that you are picking a name that is already in use as the name of another object.***

The reason for this prohibition is very simple: if you could use the name r to name this new object that you are introducing, while r is already the name of some other object that was introduced before, then you would be forcing these two objects to be the same. But there is no reason for them to be the same, so you cannot give them the same name.

Example 31. Suppose you know that Mr. Winthrop has been murdered. That means, if we use "$P(x)$" for the predicate "x murdered Mr. Winthrop". that you know that $(\exists x)P(x)$ (that is, somebody murdered Mr. Winthrop). Then you can introduce a new character into your discourse, and call this person "the murderer", or "the killer". (This is useful, because you want to be able to talk about that person, and say things such as "the murderer must have had a key so as to be able to get into Mr. Winthrop's apartment".) But you cannot call the murderer "Mrs. Winthrop", because if you do so you would be stipulating that it was Mrs. Winthrop that killed Mr. Winthrop, which could be true but you do not know that it is. □

And here is a precise statement[48] of Rule $\exists_{use}$:

Rule $\exists_{use}$

(I) If

1. $P(x)$ is a sentence,
2. the letter a is not in use as the name of anything,
3. you have proved $(\exists x)P(x)$,

then

* you can introduce a witness and call it a, so that this new object will satisfy $P(a)$

(II) In addition, if S is a set, and you have proved that $(\exists x \in S)P(x)$, then you can stipulate that $a \in S$ as well.

9.2.2 The rule for proving existential sentences (Rule $\exists_{prove}$)

This rule is very simple, and very easy to remember:

- ***to prove that there is money here, show me the money***;
- ***to prove that cows exist, show me a cow***;
- ***to prove that good students exist, show me a good student***,
- ***to prove that incorruptible politicians exist, show me an incorruptible politician***,
- ***to prove that prime numbers exist, show me a prime number,***

and so on.

Example 32. Suppose you want to prove that $(\exists x \in \mathbb{Z})x^2 + 3x = 10$.

You can say "Take $x = 2$. Then $x^2 + 3x = 10$, because $x^2 = 4$ and $3x = 6$, so $x^2 + 3x = 4 + 6 = 10$". So 2 is a witness for the sentence $(\exists x \in \mathbb{Z})x^2+3x = 10$. Then Rule $\exists_{prove}$ allows us to go to $(\exists x)x^2+3\cdot x = 10$. □

[48] In this statement, we use the same convention explained earlier: $P(a)$ is the sentence obtained from $P(x)$ by substituting a for x. For example, if $P(x)$ is the sentence "x eats grass", then P(Suzy) is the sentence "Suzy eats grass". If $P(x)$ is the sentence "$x + 3y = x^2$", then $P(a)$ is the sentence "$a + 3y = a^2$".

And here is a precise statement of the witness rule:

Rule $\exists_{prove}$

If:

1. $P(x)$ is a sentence,
2. a is a witness for $(\exists x)P(x)$ (that is, you have proved that $P(a)$),

then

* you can go to $(\exists x)P(x)$.

In addition, if S is a set, and you have proved that $a \in S$, then you can go to $(\exists x \in S)P(x)$.

In other words, ***Rule $\exists_{prove}$ says that you can prove the sentences*** $(\exists x)P(x)$ ***or*** $(\exists x \in S)P(x)$ ***by producing a witness.***

9.3 Examples of proofs involving existential sentences

9.3.1 Some simple examples

Problem 37. Consider the sentence

$$(\exists x \in \mathbb{Z})(\exists y \in \mathbb{Z})x^2 - y^2 = 17\,. \tag{9.103}$$

Is this sentence true or false? If it is true, prove it. If it is false, prove that it is false (that is, prove its negation).

SOLUTION. Sentence (9.103) is true. Here is a proof:

Take $x = 9$, $y = 8$. Then $x^2 = 81$ and $y^2 = 64$. So $x^2 - y^2 = 81 - 64 = 17$. Therefore the pair $(9, 8)$ is a witness for (9.103). By Rule $\exists_{prove}$, this proves (9.103). **Q.E.D.**

Problem 38. Consider the sentence

$$(\forall m \in \mathbb{Z})(\exists n \in \mathbb{Z})n < m\,. \tag{9.104}$$

Is this sentence true or false? If it is true, prove it. If it is false, prove that it is false (that is, prove its negation).

SOLUTION. Sentence (9.104) is true. Here is a proof.

> Let m be an arbitrary integer.
>
> We want to prove that $(\exists n \in \mathbb{Z})n < m$.
>
> For this purpose, we produce a witness. First we say who the witness is, and then we prove it works, that is, that it really is a witness.
>
> Let $\hat{n} = m - 1$.
>
> Then $\hat{n} \in \mathbb{Z}$ and $\hat{n} < m$. So the integer $\hat{n}$ is a witness for the sentence $(\exists n \in \mathbb{Z})n < m$
>
> Therefore $(\exists n \in \mathbb{Z})n < m$. [Rule $\exists_{prove}$]

Since we have proved that $(\exists n \in \mathbb{Z})n < m$ for an arbitrary integer m, we can conclude that $(\forall m \in \mathbb{Z})(\exists n \in \mathbb{Z})n < m$. [Rule $\forall_{prove}$] **Q.E.D.**

Problem 39. Consider the sentence

$$(\forall m \in \mathbb{N})(\exists n \in \mathbb{N})n < m\,. \tag{9.105}$$

Is this sentence true or false? If it is true, prove it. If it is false, prove that it is false (that is, prove its negation).

SOLUTION. Sentence (9.105) is false. Here is a proof.

> Asssume (9.105) is true.
>
> Them by Rule $\forall_{use}$ we can plug in a value for m, and the result wil be a true sentence. So we plug in $m = 1$.
>
> Them by Rule $\forall_{use}$ iimplies that $(\exists n \in \mathbb{N})n < 1$.
>
> But there is no natural number that is less than 1, so so $\sim (\exists n \in \mathbb{N})n < 1$.
>
> So we have attained a contradcition.

Therefore (9.105) is false.

Problem 40. Consider the sentence

$$(\exists n \in \mathbb{Z})(\forall m \in \mathbb{Z})n < m\,. \tag{9.106}$$

Is this sentence true or false? If it is true, prove it. If it is false, prove that it is false (that is, prove its negation).

SOLUTION. Sentence (9.106) is false. Here is a proof of its negation, that is, of

$$\sim (\exists n \in \mathbb{Z})(\forall m \in \mathbb{Z}) n < m\,. \tag{9.107}$$

We are going to prove (9.107) by contradiction .

> Assume that
>
> $$(\exists n \in \mathbb{Z})(\forall m \in \mathbb{Z}) n < m\,. \tag{9.108}$$
>
> Pick a witness for Statement (9.108), that is, an integer n for which the statement "$(\forall m \in \mathbb{Z}) n < m$" holds, and call it n_0. [Rule $\exists_{use}$]
>
> Then $n_0 \in \mathbb{Z}$ and $(\forall m \in \mathbb{Z}) n_0 < m$.
>
> Since $n_0 \in \mathbb{Z}$, we can conclude that $n_0 < n_0$. [Rule $\forall_{use}$, from $(\forall m \in \mathbb{Z}) n_0 < m$]
>
> Then $\sim n_0 = n_0$. [Trichotomy law]
>
> But $n_0 = n_0$. [Equality Axiom $(\forall x) x = x$.]

So we have proved a contradiction assuming (9.108). Hence, by the proof-by-contradiction rule, (9.108) is false, that is, (9.107) is true. **Q.E.D.**

Problem 41. For each of the following sentences,

1. Indicate whether the sentence is true or false.
2. If it is true, prove it.
3. If it is false, prove that it is false (that is, prove its negation).

$$(\forall m \in \mathbb{Z})(\exists n \in \mathbb{N}) n > m\,, \tag{9.109}$$
$$(\forall m \in \mathbb{N})(\exists n \in \mathbb{N}) n < m\,, \tag{9.110}$$
$$(\exists n \in \mathbb{N})(\forall m \in \mathbb{Z}) n < m\,, \tag{9.111}$$
$$(\exists n \in \mathbb{N})(\forall m \in \mathbb{N}) n < m\,, \tag{9.112}$$
$$(\exists n \in \mathbb{N})(\forall m \in \mathbb{N}) n \leq m\,, \tag{9.113}$$
$$(\exists x \in \mathbb{R})(\forall m \in \mathbb{N}) x < m\,. \tag{9.114}$$

9.3.2 A detailed proof of an inequality with lots of comments

Problem 42. Let C be a circle with center $(5, 1)$. Let L be the line with equation $y = x + 4$. Prove that if the radius of the circle is less than 5 then C and L do not intersect.

Solution.
Let R be the radius of C.
*COMMENT: This is very important. Every time you will have to deal repeatedly with some object—a number, a set, an equation, a statement—**give it a name**.*

Assume that $R < 5$.

We want to prove that

$$\sim (\exists x \in \mathbb{R})(\exists y \in \mathbb{R})\Big((x-5)^2 + (y-1)^2 = R^2 \wedge y = x+4 \Big)\,. \quad (9.115)$$

Assume (9.115) isn't true.

Then

$$(\exists x \in \mathbb{R})(\exists y \in \mathbb{R})\Big((x-5)^2 + (y-1)^2 = R^2 \wedge y = x+4 \Big)\,. \quad (9.116)$$

Pick witnesses for (9.116) and call them x, y.

COMMENT: Remember that after a quantified sentence ends the quantified variables become free again, so they can be re-used. That's why it is perfectly legitimate to name the witnesses x and y.

Then

$$(x-5)^2 + (y-1)^2 = R^2 \wedge y = x+4\,. \quad (9.117)$$

In particular,

$$(x-5)^2 + (y-1)^2 = R^2\,. \quad (9.118)$$

And also

$$y = x + 4\,. \quad (9.119)$$

*COMMENT: How did we go from (9.117) to (9.118) and (9.119)? It's clear, isn't it? But in a proof **every step must be justfied (or justifiable) by the rules**. So which is the rule used here? The answer is: it's the logical rule for using conjunctions, that is, Rule $\wedge_{use}$: if you have a conjunction $A \wedge B$, then you can go to A,*

and you can go to B. You may think this is a very stupid rule, but it is certainly a reasonable rule. When we went from (9.117) to (9.118) and (9.119), it seemed obvious to you, didn't it? That's because Rule $\wedge_{use}$ is an obvious rule, so obvious that you use it all the time without even noticing it. But that doesn't mean that the rule isn't there. ***It is there.*** *If you wanted to write a computer program that checks proofs and tells you whether a proof is valid, how would the program know that going from (9.117) to (9.118) and (9.119) are valid steps? You have to put that in the program. That is, you have to put Rule $\wedge_{use}$ in your program.*

Since $y = x + 4$, we can substitute $x + 4$ for y in (9.118), and get

$$(x-5)^2 + (x+4-1)^2 = R^2\,, \tag{9.120}$$

that is

$$(x-5)^2 + (x+3)^2 = R^2\,. \tag{9.121}$$

But

$$\begin{aligned}(x-5)^2 + (x+3)^2 &= x^2 - 10x + 25 + x^2 + 6x + 9\\ &= 2x^2 - 4x + 34\\ &= 2(x^2 - 2x + 17)\\ &= 2(x^2 - 2x + 1 - 1 + 17)\\ &= 2(x^2 - 2x + 1 + 16)\\ &= 2\Big((x-1)^2 + 16\Big)\\ &\geq 2 \times 16\\ &= 32\end{aligned}$$

so

$$(x-5)^2 + (x+3)^2 \geq 32\,. \tag{9.122}$$

But

$$(x-5)^2 + (x+3)^2 = R^2\,. \tag{9.123}$$

So

$$R^2 \geq 32\,. \tag{9.124}$$

COMMENT: How did we go from (9.122) and (9.123) to (9.124)? It's clear, isn't it? But in a proof ***every step must be justfied (or justifiable) by the rules****. So which is the rule used here?*

The answer is: it's the logical rule for using equality, that is, Rule $=_{use}$ (also called Rule SEE, "susbtitution of equals for equals"): if you know that an equality $s = t$—or $t = s$—holds, and you also know that some statement P involving s holds, then you can go to $P(s \to t)$, where $P(s \to t)$ is the statemenet pbtained from P by substituting t for s in P. You may think this is a very stupid rule, but it is certainly a reasonable rule. When we went from (9.122) and (9.123) to (9.124), it seemed obvious to you, didn't it? That's because Rule SEE is an obvious rule, so obvious that you use it all the time without even noticing it. But that doesn't mean that the rule isn't there. ***It is there.***

If you wanted to write a computer program that checks proofs and tells you whether a proof is valid, how would the program know that going from (9.122) and (9.123) to (9.124) is a valid step? You have to put that in the program. That is, you have to put Rule SEE in your program.

But we are assuming that $R < 5$, and then $R^2 < 25$.

COMMENT: That's because R is positive. If all you know about was that R is a real number and $R < 5$, then R could be -10, in which case it would not follow that $R^2 > 25$. But in our case R is the radius of a circle, so $R > 0$, and the conclusion that $R < 25$ follows.

So $\sim R^2 \geq 32$. But $R^2 \geq 32$. So we have proved a contradiction.

COMMENT: The contradiction is the statement "$R^2 \geq 32 \wedge \sim R^2 \geq 32$". This is a contradiction because it is fo the form $Q \wedge \sim Q$, where Q is the statement "$R^2 \geq 32$".

So (9.115) is proved. **Q.E.D.**

9.3.3 The same proof without the comments

Proof. Let R be the radius of C.

Assume that $R < 5$.

We want to prove that

$$\sim (\exists x \in \mathbb{R})(\exists y \in \mathbb{R})\Big((x-5)^2 + (y-1)^2 = R^2 \wedge y = x + 4 \Big). \quad (9.125)$$

Assume (9.125) isn't true. Then

$$(\exists x \in \mathbb{R})(\exists y \in \mathbb{R})\Big((x-5)^2+(y-1)^2 = R^2 \wedge y = x+4 \Big) . \quad (9.126)$$

Pick witnesses for (9.126) and call them x, y.

Then $(x-5)^2 + (y-1)^2 = R^2 \wedge y = x+4$, so in particular,

$$(x-5)^2 + (y-1)^2 = R^2 . \quad (9.127)$$

Since $y = x+4$, we can substitute $x+4$ for y in (9.127), and get $(x-5)^2 + (x+4-1)^2 = R^2$, that is

$$(x-5)^2 + (x+3)^2 = R^2 . \quad (9.128)$$

But

$$\begin{aligned} (x-5)^2 + (x+3)^2 &= x^2 - 10x + 25 + x^2 + 6x + 9 \\ &= 2x^2 - 4x + 34 \\ &= 2(x^2 - 2x + 17) \\ &= 2(x^2 - 2x + 1 - 1 + 17) \\ &= 2(x^2 - 2x + 1 + 16) \\ &= 2\Big((x-1)^2 + 16\Big) \\ &\geq 2 \times 16 \\ &= 32 \end{aligned}$$

so

$$(x-5)^2 + (x+3)^2 \geq 32 . \quad (9.129)$$

But $(x-5)^2 + (x+3)^2 = R^2$, so $R^2 \geq 32$.

But we are assuming that $R < 5$, and then $R^2 < 25$.

So $\sim R^2 \geq 32$. But $R^2 \geq 32$. So we have proved a contradiction.

So (9.125) is proved. **Q.E.D.**

9.4 Existence and uniqueness

Suppose $P(x)$ is a one-variable predicate. We write

$$(\exists! x)P(x)$$

for "there exists a unique x such that $P(x)$."

This means "there is one and only one x such that $P(x)$".

The precise meaning of this is that

1. there exists an x such that $P(x)$,

and

2. if x_1, x_2 are such that $P(x_1) \wedge P(x_2)$, then $x_1 = x_2$.

In formal language:

$$(\exists! x)P(x) \iff \Big((\exists x)P(x) \wedge \Big((\forall x_1)(\forall x_2)(P(x_1) \wedge P(x_2)) \implies x_1 = x_2\Big)\Big).$$

It follows that, in order to prove that there exists a unique x such that $P(x)$, you must prove two things:

Existence: There exists x such that $P(x)$,

Uniqueness: Any two x's that satisfy $P(x)$ must be equal.

That is:

To prove

$$(\exists! x)P(x)$$

it suffices to prove

$$(\exists x)P(x) \tag{9.130}$$

and

$$(\forall x_1)(\forall x_2)\Big((P(x_1) \wedge P(x_2)) \implies x_1 = x_2\Big). \tag{9.131}$$

(Formula (9.130) is the existence assertion, and Formula (9.131) is the uniqueness assertion.)

Example 33. "I have one and only one mother" means:

- I have a mother,

and

- Any two people who are my mother must be the same person. (That is: if u is my mother and v is my mother than $u = v$.) □

9.4.1 Examples of proofs of existence and uniqueness

Problem 43. Prove that there exists a unique natural number n such that $n^3 = 2n - 1$.

Solution. We want to prove that

$$(\exists! n \in \mathbb{N}) n^3 = 2n - 1\,.$$

First let us prove existence. We have to prove that $(\exists n \in \mathbb{N}) n^3 = 2n - 1$. To prove this, we exhibit a witness: we take $n = 1$. Then n is a natural number, and $n^3 = 2n - 1$. So $(\exists n \in \mathbb{N}) n^3 = 2n - 1$.

Next we prove uniqueness. We have to prove that if u, v are natural numbers such that $u^3 = 2u - 1$ and $v^3 = 2v - 1$, then it follows that $u = v$.

So let u, v be natural numbers such that $u^3 = 2u - 1$ and $v^3 = 2v - 1$. We want to prove that $u = v$.

Since $u^3 = 2u - 1$ and $v^3 = 2v - 1$, we have

$$\begin{aligned} u^3 - v^3 &= 2u - 1 - (2v - 1) \\ &= 2u - 2v \\ &= 2(u - v)\,, \end{aligned}$$

so

$$u^3 - v^3 - 2(u - v) = 0\,.$$

But it is easy to verify that

$$u^3 - v^3 = (u - v)(u^2 + uv + v^2)\,.$$

(If you do not believe this, just multiply out the right-hand side and you will find that the result equals $u^3 = v^3$.) Hence

$$\begin{aligned} 0 &= u^3 - v^3 - 2(u - v) \\ &= (u - v)(u^2 + uv + v^2) - 2(u - v) \\ &= (u - v)(u^2 + uv + v^2 - 2)\,. \end{aligned}$$

We know that if a product of two real numbers is zero then one of the numbers must be zero. Hence

$$u - v = 0 \quad \text{or} \quad u^2 + uv + v^2 - 2\,.$$

But $u^2 + uv + v^2 - 2$ cannot be equal to zero, because u^2, uv and v^2 are natural numbers, so each of them is gretar than or equal to 1, and then $u^2 + uv + v^2 \geq 3$, so $u^2 + uv + v^2 - 2 \geq 1$, and then $u^2 + uv + v^2 - 2 \neq 0$. Therefore $u - v = 0$, so $u = v$, and our proof of uniqueness is complete.

Problem 44. Prove that there exists a unique real number x such that

$$x^7 + 3x^5 + 23x = 6\,.$$

You are allowed to use everything you know from Calculus. □

10 A summary of Logic

10.1 Terms and sentences

10.1.1 Nouns and noun phrases in English

- According to *Wikipedia*, a noun is "a word that functions as the name of some specific thing or set of things, such as living creatures, objects, places, actions, qualities, states of existence, or ideas".
- A noun phrase "is a phrase that has a noun (or indefinite pronoun) as its head or performs the same grammatical function as such a phrase".

So, for example, here is a list of noun phrases:

1. George Washington,
2. the first president of the United States,
3. the man who succeeded George Washington as prsdient of the United States,
4. this table,
5. the table,
6. the table that I bought yesterday,
7. the table that I bought yesterday at Walmart's and then brought home in a truck that I had borrowed from my very good friend Alice,
8. I,
9. you,
10. she,
11. he,
12. the news,
13. the number five,
14. the number that results from adding two plus three,
15. the product of two and three,
16. the number that results from adding two plus three and then multiplying the result by four,
17. the number that results from adding two plus three, multiplying the result by four, and then dividing the result of the multiplication by the product of six times seven,
18. the sum of the cubes of all the natural numbers from eight to forty-seven.

10.1.2 The "use-mention" distinction

Consider the following two sentences:

> Clarabelle is a cow.
>
> "Clarabelle" is a ten-letter word.

The first sentence talks about an animal and makes an assertion about that animal: it tells us that that animal is a cow.

The second sentence does not talk about an animal. It does not say anything about the animal called Clarabelle. It makes an assertion about a ***word***: it tells us that the word "Clarabelle" has ten letters.

The first sentence talks about the animal, so it ***mentions*** Clarabelle, And, in order to mention (i.e., talk about) Clarabelle it ***uses*** the word "Clarabelle".

The second sentence talks about the word, so it ***mentions*** the word "Clarabelle".

So the first sentence ***uses*** the word "Clarabelle" and the second sentence ***mentions*** it.

The distinction between use and mention is very important, and it is useful to understand it in order to avoid making mistakes in writing that sometimes might be confusing to the reader.

Let us be precise: word and groups of words are things, and like all other things words and groups of words can be given names. ***the name of a word or group of words is the word or groups of words enclosed in quotation marks***.

So, for example, the following are correct statements:

- Clarabelle is an animal.
- The name of Clarabelle is "Clarabelle".
- "Clarabelle" is a word.
- "Clarabelle" is a French name, not a cow.
- Clarabelle is a cow, not a French name;
- Clarabelle eats grass.
- "Clarabelle" does not eat grass.
- The name of the word "Clarabelle" is " "Clarabelle" ".
- The name of the first president of the United States was "George Washington". (If you had written instead "the name of the first president of the United States was George Washington", then, since George Washington was a general, it would follow that the name of the first president of the United States was a general, which is quite ridiculous, since a name is a word or group of words, and cannot be a general.)

- The name of George Washington is[49] "George Washington".
- If I say "the name of that cow over there is Clarabelle", then I am saying among other things that the name of that cow over there is a cow, which is not what I probably want to say[50]. I probably want to say that that the name of that cow over there is the word "Clarabelle", so I must say: the name of that cow over there is "Clarabelle".

When you say something about Clarabelle, the cow, you ***use*** the word "Clarabelle" to talk about the cow, and by doing so you ***mention*** (i.e., talk about) the cow.

When you say something about "Clarabelle", the word, you ***mention*** (i.e., talk about) the word "Clarabelle", but you are not using the word to talk about the animal.

When you ***use*** a word or group of words to talk about the thing that the word stands for, you do not enclose the word or group of words in quotation marks.
When you ***mention*** a word or groups of words (i.e.,. talk about the word or gourp of words), you ***must*** enclose the word or group of words in quotation marks.

When writing mathematics, it is important to keep the distinction between use and mention, by using quotation marks when appropriate.

For example,

- we can write

  $$\text{I will prove that } 2+2=4$$

  in the same was as we would write

  Alice said that she likes coffee.

- but we should not write

  $$\text{I will prove } 2+2=4$$

[49] So, strictly speaking, it is wrong to write: "my name is Alexander Hamilton", or "my name is Asher Lev", or "my name is Eminem". One should write "my name is "Alexander Hamilton" ", or "my name is "Asher Lev" ", or "my name is "Eminem" ". But this mistake is so common that nobody pays attention to it.

[50] For example: Clarabelle eats grass. So, if the name of that cow over there is Clarabelle, it follows that the name of that cow over there eats grass. And this is nonsense, of course: *cows* eat grass, *names* do not.

in the same was as we would not write

Alice said I like coffee.

We must write

Alice said "I like coffee".

and

I will prove "$2 + 2 = 4$",

or

I will prove that "$2 + 2 = 4$" is true,

or

I will prove that the sentence "$2 + 2 = 4$" is true,

or

I will prove the sentence "$2 + 2 = 4$".

10.1.3 Terms in mathematical language

The noun phrases that we use in formal mathematical language are called ***terms***.

So a ***term*** is an expression that is the name of a thing. For example[51] the terms "four", "4", "two plus two", "$2 + 2$" , "three plus one", $3 + 1$" all have the same value, namely, the number 4.

And usually mathematical terms are written with ***formulas***, that is, very concise expressions using special symbols. For example,

- instead of "the number five", we write "5";
- instead of "the number that results from adding two plus three", we write "$2 + 3$";
- instead of "the product of two and three", we write "2×3";
- instead of "the number that results from adding two plus three and then multiplying the result by four", we write "$(2 + 3) \times 4$";

[51] Notice the use of the quotation marks, in keeping with the *use vs. mention* distinction explained in subsection 10.1.2. We can say correctly that 4 is a number, that $2 + 2$ is a number, that the term "4" has the value 4, that the term "$2 + 2$" has the value 4, that $2 + 2 = 4$ (meaning that both terms "$2 + 2$" and "4" have the same value). But it would be incorrect to write "4" = "$2 + 2$" because this says that the two terms "4" and "$2 + 2$" are the same, which is not true. (For example, the term "4" consists of just one character, whereas the term "$2+2$" consists of three characters, so they are manifestly not the same.)

- instead of "the number that results from adding two plus three, multiplying the result by four, and then dividing the result by the sum of twenty-three and the product of six times seven"; we write "$\frac{(2+3)\times 4}{23+6\times 7}$";
- instead of "the sum of the cubes of all the natural numbers from five to ten" we write "$\sum_{i=5}^{10}$".

10.1.4 Examples of terms and sentences

Example 34. The following expressions are terms:

- New York City;
- Mount Everest;
- the table;
- the student who asked why an implication is true when the premise is false;
- 2,
- $2 + 2$,
- $2 + x$,
- $x + y$,
- $(x + y)^2 + 3x + 5$,
- $\sum_{k=1}^{n}(k^3 + 1)$

But the following expressions are sentences, not terms:

- $2 + 2 = 4$,
- $2 + x = 4$,
- $x + y = 0$,
- $x + y = 0$,
- $x + y > 0$,
- $(\exists y \in \mathbb{R})x + y = 0$,
- $(\exists y \in \mathbb{R})x + y < 0$,

- $(\forall y \in \mathbb{R})x + y = 0$,
- $(\forall y \in \mathbb{R})(\exists z \in \mathbb{R})y + z = x$.
- $(\forall x \in \mathbb{R})x.0 = 0$,
- $(\forall x \in \mathbb{R})(\exists y \in \mathbb{R})x + y > 0$,
- $(\forall x \in \mathbb{R})x^2 \geq 0$,
- $\sum_{x=1}^{5} x^2 = y$.
- $(\forall x \in \mathbb{R})x.0 = 0$,
- $\sum_{x=1} 5x^2 = y$.

10.1.5 The value of a term

A term has a ***value***, which is the thing that the term stands for. For example,

- the value of the term "5" is the natural number 5, and we indicate this by writing "$5 = 5$";
- the value of the term "$2 + 3$" is the natural number 5, and we indicate this by writing "$2 + 3 = 5$";
- the value of the term "2×3" is the natural number 6, and we indicate this by writing "$2 \times 3 = 6''$;
- the value of the term "$(2 + 3) \times 4$" is the natural number 20, and we indicate this by writing "$(2 + 3) \times 4 = 20$";
- the value of the term "$\frac{(2+3)\times 4}{23+6\times 7}$" is the rational number (i.e., fraction) $\frac{20}{65}$, and we indicate this by writing "$\frac{(2+3)\times 4}{23+6\times 7} = \frac{20}{65}$"; furthermore, the number $\frac{20}{65}$ is the same as the number $\frac{4}{13}$, so we could also written "$\frac{(2+3)\times 4}{23+6\times 7} = \frac{4}{13}$";
- the value of the term "$\sum_{i=5}^{10} i^3$" is the natural number $2,955$, and we indicate this by writing the equality "$\sum_{i=5}^{10} i^3 = 2,955$".

The values of terms can be all kinds of mathematical objects. Since the mathematical objects that you are most familiar with are numbers (natural numbers, integers, real numbers, complex numbers, etc.), you are probably used to terms whose values are numbers. But there are millions of other

kinds of mathematical objects, and we can write terms with values of any of those kinds.

For example:

- The expression
$$\begin{bmatrix} 1 & 2 \\ -1 & 3 \end{bmatrix} + \begin{bmatrix} 2 & 2 \\ 3 & 1 \end{bmatrix}$$
is a term whose value is a 2 by 2 matrix. The actual value of the term is the matrix $\begin{bmatrix} 3 & 4 \\ 2 & 4 \end{bmatrix}$, and we indicate this by writing:
$$\begin{bmatrix} 1 & 2 \\ -1 & 3 \end{bmatrix} + \begin{bmatrix} 2 & 2 \\ 3 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 4 \\ 2 & 4 \end{bmatrix}$$

- ***Functions***[52] can be added (in some cases), and composed (in some cases). If f and g are functions then the name of the sum of f and g is "$f + g$", the name of the product of f and g is "$f \cdot g$", and the name of the composite "g followed by f" is "$f \circ g$". So, for example, if f, g, h are three functions, then the expression "$((f + g) \cdot g) \circ h$" is a term whose value is a function.

- ***Sets***[53] can be combined in various ways. For example, of A and B are sets, then we can form the sets $A \cup B$ (the *union* of A and B), $A \cap B$ (the *intersection* of A and B), $A \times B$ (the *Cartesian product* of A and B). Then the value of the term "$(\mathbb{R} \times \mathbb{Z}) \cap (\mathbb{Z} \times \mathbb{R})$" is a set, namely, the intersection of the Cartesian product of $\mathbb{R}$ and $\mathbb{Z}$ with the Cartesian product of $\mathbb{Z}$ and $\mathbb{R}$.

10.1.6 Terms as instructions for a computation, i.e., as programs

You should think of a mathematical term as a computing device that executes a program, i.e., follows with a list of instructions for computing a result, called the ***value***. For example,

- the term "$2+3$" is a device that executes the following program: "add the number 3 to the number 2 and write down the result";
- the term "2×3" is a device that executes the following program: "multiply the number 2 by the number 3 and write down the result";

[52] We will talk about functions later in the course.

[53] We will discuss sets in detail later later in the course.

- the term "$(2+3)\times 4$" is a device that executes the following program: "add the number 3 to the number 2, multiply the result by the number 4, and write down the result";
- the term "$\frac{(2+3)\times 4}{23+6\times 7}$"; is a device that executes the following program: "add the number 2 to the number 3, multiply the result by the number 4, and divide the result by the number you get when you add to the number twenty-three the product of six times seven, and write down the result";
- the term "$\sum_{i=5}^{10}(i^3+3i^2+5)$" is a device that executes the following program:

  1. Look at all the natural numbers from 5 to 10.
  2. For each such natural number, do the following:
     (a) Call the number i.
     (b) Compute $(i^3+3i^2+5)^2$.
  3. Add up all the results of the computation of $(i^3+3i^2+5)^2$ for all values of i.
  4. Write the result of this sum as the final result of the computation.

10.1.7 Letter variables in terms

A term can contain ***variables***, i.e. symbols that are not the names of definite objects, but could be used to stand for different objects.

For example: The term "$x+3$", contains the letter variable x; it corresponds to the program: "add 3 to x and write down the result". When asked to compute $x+3$, the term does not know what to do, because it does not know who x is. But if you give x a specific value, for example by saying "Let $x=2$", then the term knows what to do: it gives x the value 2, and becomes the term $2+3$, which then know what to, and compues the value 5.

In other words: if a term t contains a variable x, then it is possible to give a value to the variable, and the term then can compute a value.

You should think of a variable as a "slot" that can be filled by plugging in a value. For example, the term "$x+3$" consists of (1) a slot that can be filled in with a number; (2) the + sign, (3) the number 3.

A term may contain several variables. For example, the term

$$(x+y+3x^2)y+y^2(z^2+3xz)+ye^x$$

contains the variables x, y, and z. The term has 10 slots. You can give a value to each of the three variables. The term then instructs us to fill in

the first, third, seventh and tenth slots (the "x-slots") with the value we have chosen for x, the second, fourth, fifth and ninth slots (the "y-slots") with the value we have chosen for y, and and the sixth and eight slots (the "z-slots") with the value we have chosen for z. We can do this by writing, for example:

$$\text{Let } x = 3\,,\ y = -1\,,\ z = 4\,.$$

$$\text{Then } (x + y + 3x^2)y + y^2(z^2 + 3xz) + ye^x = 33 - e^3\,.$$

or

$$\Big((x + y + 3x^2)y + y^2(z^2 + 3xz) + ye^x\Big)_{x=3,y=-1,z=4} = 33 - e^3\,.$$

In a term t, a letter variable that has no value declared within the term and represents a slot that can be filled in by giving it values is called a free variable, or open variable.

10.1.8 Bound (dummy, closed) variables in terms

One of the main purposes of writing terms and sentences in formal language, with symbols, rather than phrases with lots of words, is to be able to say things much more ***concisely***[54]. (This is quite clear: "$2 + 2 = 4$" is much shorter than "two plus two equals four". And try to say "$(a + b)^2 = a^2 + 2ab + b^2$" with words, rather than symbols, and you'll se how much longer it gets.)

[54]This is ***not*** the only purpose. ***Another purpose*** is ***precision***: for example, if I say "two plus three times five", then this is ambiguous, because it could mean "two plus the product of three and five", or "the sum of three plus two, multiplied by five". In formal language, we write "$(2 + 3) \times 4$" or "$2 + (3 \times 5)$", and each of these two expressions has a clear and precise meaning. The ambigüity has disappeared. Furthermore, we agreed on the convention that when a product such as 3×4 is combined with another term by a "+" the parentheses surrounding the product can be omitted. So when we think we ought to write "$2 + (3 \times 4)$" we write instead "$2 + 3 \times 4$", and it is completely clear what that means, because if we had wanted to say "$(2 + 3) \times 4$" we would have had to enclose "$2 + 3$" in parenthees. A ***third purpose*** is ***universality***: to say "two plus two equals four" in Spanish you have to say "dos más dos es igual a cuatro", and in French you have to say "deux plus deux égale quatre". But in formal mathematical language you write "$2 + 2 = 4$", and this is understood by everybody, whether they speak English or Spanish or French or Chinese or any other language.

Further conciseness can be achieved by using letters to stand for expressions that appear repeatedly in a term and are very long. For example, the term

$$32 + 5\frac{1+\sqrt{5}}{2} - 23\left(\frac{1+\sqrt{5}}{2}\right)^2 + +7\left(\frac{1+\sqrt{5}}{2}\right)^3 + 19\left(\frac{1+\sqrt{5}}{2}\right)^4 \tag{10.132}$$

can be written as

$$(32 + 5x - 23x^2 + 7x^3 + x^4)_{x=\frac{1+\sqrt{5}}{2}} \ , \tag{10.133}$$

which we read as the computing instruction "give x the value $\frac{1+\sqrt{5}}{2}$, then compute $32 + 5x - 23x^2 + 7x^3 + x^4$, and write down the result".

Another example is the term

$$1^3 + 2^3 + 3^3 + 4^3 + 5^3 + 6^3 + 7^3 + 8^3 + 9^3 + 10^3 \, , \tag{10.134}$$

that can be written as

$$\sum_{k=1}^{10} k^3 \, , \tag{10.135}$$

which we read as the computing instruction

1. Look at all the natural number values between 1 and 10.
2. For each such value, do the following:
   (a) Call your number k.
   (b) Compute k^3.
3. Add up the results of these computations, for all natural numbers between 1 and 10.

As you can see, the term (10.133) is much shorter than the term (10.132), and the term (10.135) is much shorter than the term (10.134). And the difference becomes even more dramatic if consider very long terms, in which there is a computation that is repeated over and over. For example, suppose you want to talk about the sum of the cubes of the first $10,000$ natural numbers: using letters, we can write

$$\sum_{k=1}^{10{,}000} k^3 \, . \tag{10.136}$$

If you wanted to write this without using the $\sum$ notation, you would have to write a sum of $10,000$ terms, which would of course be enormously long[55].

It is clear that:

- In (10.133), we could have used any other letter instead of x, and the resulting term would execute exactly the same computation. So, for example, if we had written

$$(32 + 5u - 23u^2 + 7u^3 + u^4)_{u=\frac{1+\sqrt{5}}{2}}, \tag{10.137}$$

  this would describe exactly the same computationas (10.137), and the term would have exactly the same value and (10.133).

- Similarly, in (10.135), we could have used any other letter instead of k, and the resulting term would execute exactly the same computation. So, for example, if we had written

$$\sum_{i=1}^{10} i^3 \text{ , or } \sum_{j=1}^{10} j^3 \text{ , or } \sum_{x=1}^{10} x^3 \text{ , or } \sum_{\alpha=1}^{10} \alpha^3 \text{ , or } \sum_{\diamond=1}^{10} \diamond^3 ,$$

  the resulting term would correspond to exactly the same computation and have the same value[56].

- Actually, in the term "$\sum_{k=1}^{10} k^3$" the letter k "isn't there", in the sense that we could describe the term without ever mentioning "k". For example, I could ask you to compute the value of this term without mentioning k: by saying "compute the sum of the cubes of all the natural numbers from 1 to 10", and you would know exactly what to do.

- A similar situation arises for the term

$$(32 + 5x - 23x^2 + 7x^3 + x^4)_{x=\frac{1+\sqrt{5}}{2}} .$$

  The letter x "isn't there", in the sense that we could describe the term without ever mentioning "x". For examle, I could ask you to compute the value of this term without mentioning x: by saying

[55] Can you figure out what the value of this sum is? The answer is: $5,001,000,050,000,000$. Can you figure this out without having to compute $10,000$ cubes and then add them? Later in the course we will see how to figure this out and get the answer fairly fast.

[56] Using "x" here is not something one would normally do, because mathematicians usually prefer to use "x" for real numbers rather than natural numbers; but it is not forbidden to use x for a natural number.

a. Add the following five numbers:
   1. the number 32,
   2. the number $5 \times \frac{1+\sqrt{5}}{2}$,
   3. the number $(-23) \times \left(\frac{1+\sqrt{5}}{2}\right)^2$,
   4. the number $7 \times \left(\frac{1+\sqrt{5}}{2}\right)^3$,
   5. the number $\left(\frac{1+\sqrt{5}}{2}\right)^4$.

b. Then write down the result.

10.1.9 What is a dummy (free, open) variable?

In a term t, a letter variable whose values are generated within the term itself, so that we do not need to ask the outside world what the value of that variable is in order to be able to compute the value of the term, is called a bound variable, or closed variable, or dummy variable.

The three clear signs that a variable is dummy

The following are three obvious signs that a variable in an expression T is a dummy variable:

(I) It is possible to substitute for the variable any other letter without changing the value of the expression. Indeed, the terms

$$\sum_{j=1}^{N} j^2\,, \tag{10.138}$$

$$\sum_{x=1}^{N} x^2\,, \tag{10.139}$$

$$\sum_{q=1}^{N} q^2\,, \tag{10.140}$$

all have the same value as "$\sum_{k=1}^{N} k^2$", and this is a sign that in the term "$\sum_{k=1}^{N} k^2$" the letter k is a dummy variable.

(II) If you ask somebody to execute the computation described by the expression T then this person does not need to be told what the value of the variable is, because the computation itself generates the value or values it needs for the variable. For example, if I ask you to compute the value of "$\sum_{k=1}^{N} k^2$", then you have to do this: you give k all natural number values between 1 and N, for each such value you compute its square, and then you add all the results. In order to be able to do this, you have to ask "who is N?", but ***you do not have to ask "who is k?"***, because you yourself, in the process of doing the computation, are going to generate the values of k. This is a second sign that in the term "$\sum_{k=1}^{N} k^2$" the letter k is a dummy variable.

(III) The expression T is equal or equivalent to another expression not involving the variable at all. For example, "$\sum_{k=1}^{N} k^2$" is equal to $\frac{(2N+1)N(N+1)}{6}$, an expression that does not contain k. And this is a third sign that in the term "$\sum_{k=1}^{N} k^2$" the letter k is a dummy variable.

10.1.10 Other examples of dummy variables in terms

The two examples of dummy variables that you probably know from previous courses are those occurring in expressions such as "$\sum_{k=a}^{b} \ldots$" and "$\prod_{k=a}^{b} \ldots$".

For example, the term "$\sum_{k=1}^{5} k^2$" executes the following computation: look at all the natural numbers from 1 to 45, for each such number compute

its square, add all the results and write down the sum".

And the term "$\prod_{k=1}^{5}(k+1)^2$" executes the following computation: look at all the natural numbers from 1 to 45, for each such number compute the square of the sum of one plus the number, multiply all the results and write down the sum".

In both cases, the letter k is a dummy variable. You do not need to ask "who is k'?" is order to carry out the computation. If you are asked to compute $\sum_{k=1}^{25} k^3$ or $\prod_{k=1}^{25}(k+1)$, then you do not have to ask "who is k?". You yourself generate all the values of k from 1 to 25 and for each value compute something (k^3 in the first case, $k+1$ oin the second case), and then do something with the results (add them all up in the first case, multiply them in the second case).

Variables of integration. Another important example of a dummy variable is a ***variable of integration.*** If I ask you to tell me what the value of the integral

$$\int_a^b x^2\,dx$$

is, you have to ask me "who are a and b" but you don't have to ask "who is x?". That's because x is a dummy variable. This is precisely the second of the three "signs that a variable is a dummy variable".

Let us look at the first sign: "It is possible to substitute the letter for any other letter, without changing the value of the term". This is indeed true: if, instead of "$\int_a^b (x+1)^2\,dx$" I write "$\int_a^b (y+1)^2\,dy$", or "$\int_a^b (u+1)^2\,du$", or "$\int_a^b (q+1)^2\,dq$", or "$\int_a^b (\alpha+1)^2\,d\alpha$", then all those integrals are exactly the same.

Finally, let us look at the third sign: "The term T is equal to another expression not involving the variable at all." And this is indeed true: the integral $\int_a^b (x+1)^2\,dx$ is equal to $\frac{(b+1)^3}{3}-\frac{(a+1)^3}{3}$. And the expression "$\frac{(b+1)^3}{3}-\frac{(a+1)^3}{3}$" does not contain x at all.

So, clearly, ***in the term "$\int_a^b x^2\,dx$", the variable x is a dummy variable.*** This means that ***the integral $\int_a^b x^2 dx$ does not depend on x; it depends on a and b but not on x, in the sense that if you want me to give you a specific value for the term then you have to tell me who a and b are, but not who x is.***

Variables in the definition of a function We will be discussing ***functions*** later. But let me tell you the basic facts:

- A ***function*** assigns to each object x belonging to a certain set S

another object, called the ***value*** of the function at x. If f is a name of the function, then we use $f(x)$ to denote the value of f at x.

- The set S is called the ***domain*** of the function.
- In order to introduce and describe a function f, we can do this: we explain how, for each member of the domain S, the value of f at this member is computed or determined. To do this, we write things like

$$\text{the function } \mathbb{R} \ni x \mapsto 3(x+1)^4 + e^x\,, \tag{10.141}$$

which says that (a) the domain of the function is the set $\mathbb{R}$, (b) for each member of the domain $\mathbb{R}$, if we call that member x, then the value $f(x)$ is the number $3(x+1)^4 + e^x$.

Notice that "x" is a name that we introduce in order to explain how to compute $f(x)$. We could equally well have written:

$$\text{the function } \mathbb{R} \ni u \mapsto 3(u+1)^4 + e^u,$$

and that would be exactly the same function.

So ***the variable x in a function definition such as (10.141) is dummy.***

Remark 10. An expression such as "the function $\mathbb{R} \ni x \mapsto x^2$" is a term. Its value is a function, namely, the function that takes each real number and squares it. The term "the function $\mathbb{R} \ni x \mapsto x^2$" contains the variable x but, as we just explained, x is a dummy variable, because (a) the term is equal to another term that does nto contain x at all (namely, the term "the function that for each real number produces as value the square of the number", or "the function that takes each real number and suares it"); (b) if we substitute another letter for x we get the same function. (For example, "the function $\mathbb{R} \ni u \mapsto u^2$" is exactly the same function: it's "the function that takes each real number and squares it".)

The expressions "the function $\mathbb{R} \ni x \mapsto x^2$" and "the function $\mathbb{R} \ni x \mapsto (x+1)^2 - 2x - 1$" are terms. Each of these terms has a value, which is a function. Furthermore, those two functions are the same function, because for every real number x the numbers x^2 and $(x+1)^2 - 2x - 1$ are equal. So the terms "the function $\mathbb{R} \ni x \mapsto x^2$" and "the function $\mathbb{R} \ni x \mapsto (x+1)^2 - 2x - 1$" have the same value, and we can assert that

$$\text{The function } \mathbb{R} \ni x \mapsto x^2 \;= \text{the function } \mathbb{R} \ni x \mapsto (x+1)^2 - 2x - 1.$$

NOTE: It would be incorrect to write

"The function $\mathbb{R} \ni x \mapsto x^2$" = "the function $\mathbb{R} \ni x \mapsto (x+1)^2 - 2x - 1$",

because this asserts that the two terms are the same, which is manifestly not the case. If this is not clear tou consider the following examples:

Bill Clinton = William Jefferson Clinton

and

"Bill Clinton" = "William Jefferson Clinton".

The first one says that that the values of the terms "Bill Clinton" and "William Jefferson Clinton" are the same, i.e., that the person whose name is "Bill Clinton" is the same as the person whose name is "William Jefferson Clinton".

But the seond one says that the names "Bill Clinton" and "William Jefferson Clinton" are the same, i.e. that they conist of exactly the same letters in the same order. nd this is clearly false. (For example, the name "William Jefferson Clinton" contains a W and a J, but the name "Bill Clinton" does not. □

10.1.11 Bound (dummy, closed) variables in sentences

Sentences are very similar to terms. Like terns, sentences have ***values***. The one big difference between terms and sentences is that ***the value of a term is a thing and the value of a sentence is a truth value, i.e., "true" or "false".***

Example 35.

1. The expression "$2 + 2$" is a term. Its value is the number 4.

2. The expressions "$2 + 2 = 4$" and "$2 + 2 = 5$" are sentences. They are both propositions, because they contain no open variables. So they both have a truth value. The value of "$2+2 = 4$" is "true". The value of "$2 + 2 = 5$" is "false".

3. The expression "$(\forall n \in \mathbb{Z})(2|n \implies 4|n^2)$" is a sentence. It contains the variable n . But this variable is dummy (closed, bound), because it satisfies all the three signs for dummy variables that we saw before:

(I) *It is possible to substitute for the letter n any other letter, without changing the value of the expression.* Indeed, the sentences

$$(\forall m \in \mathbb{Z})(2|m \Longrightarrow 4|m^2)\,, \tag{10.142}$$
$$(\forall q \in \mathbb{Z})(2|q \Longrightarrow 4|q^2)\,, \tag{10.143}$$
$$(\forall u \in \mathbb{Z})(2|u \Longrightarrow 4|u^2)\,, \tag{10.144}$$

are all equivalent to "$(\forall n \in \mathbb{Z})(2|n \Longrightarrow 4|n^2)$".

(II) *If you ask somebody to execute the computation described by this sentence, then this person does not need to be told what the value of the variable n, is, because the computation itself generates the value or values it needs for the variable.* (Indeed, to execute the computation described by the sentence "$(\forall n \in \mathbb{Z})(2|n \Longrightarrow 4|n^2)$", the person doing the computation has to do the following:

(a) look at all the integers, and for each integer do the following:
   i. call the integer n,
   ii. determine if "$2|n \implies 4|n^2$ is true,

(b) then look at all the results of the computations, for all the integers. If the are all "true" write "true". If one of the results is "false", write 'false".

The key point here is that ***the person doing the computation does need to ask "who is n?" because they themselves will generate the values of n to be looked at.***

(III) *The sentence "$(\forall n \in \mathbb{Z})(2|n \Longrightarrow 4|n^2)$" is equivalent to another sentence not involving the variable n at all.* Indeed: the sentence "$(\forall n \in \mathbb{Z})(2|n \Longrightarrow 4|n^2)$" is equivalent to "if an integer is even then its square is divisible by 4".

Since the only variable that occurs in this sentence is bound, the sentence contains no open variables. So it is a proposition. It has a definite truth value, which turns oiut to be "true".

4. The expression "$(\forall n \in \mathbb{Z})(a|n \Longrightarrow b|n^2)$" is a sentence. It contains the variables n, a, b . But b variable is dummy (closed, bound), because it satisfies all the three signs for dummy variables that we saw before:

(I) *It is possible to substitute for the letter n any other letter, without*

changing the value of the expression. Indeed, the sentences

$$(\forall m \in \mathbb{Z})(a|m \Longrightarrow b|m^2)\,, \tag{10.145}$$
$$(\forall q \in \mathbb{Z})(a|q \Longrightarrow b|q^2)\,, \tag{10.146}$$
$$(\forall u \in \mathbb{Z})(a|u \Longrightarrow b|u^2)\,, \tag{10.147}$$

are all equivalent to "$(\forall n \in \mathbb{Z})(a|n \Longrightarrow b|n^2)$".

(II) *If you ask somebody to execute the computation described by this sentence, then this person does not need to be told what the value of the variable n, is, because the computation itself generates the value or values it needs for the variable.* (Indeed, to execute the computation described by the sentence "$(\forall n \in \mathbb{Z})(a|n \Longrightarrow b|n^2)$", the person doing the computation has to do the following:

(a) look at all the integers, and for each integer do the following:
   i. call the integer n,
   ii. determine if "$a|n \implies b|n^2$ is true,

(b) then look at all the results of the computations, for all the integers. If the are all "true" write "true". If one of the results is "false", write 'false".

The key point here is that ***the persons doing the computation does need to ask "who is n?" because they themselves will generate the values of n to be looked at.***

This is to be contrasted with a and b. The person doing the computation cannot do anything without asking first "'who are a and b?". So a and b are free variables.

(III) *The sentence "$(\forall n \in \mathbb{Z})(a|n \Longrightarrow b|n^2)$" is equivalent to another sentence not involving the variable n at all.* Indeed: "$(\forall n \in \mathbb{Z})(a|n \implies b|n^2)$" is equivalent to "if an integer is divisible a, then its square is divisible by b". And you can see that this sentence contains a and b but not n.

5. The expression

$$p \in \mathbb{Z} \wedge p > 1 \wedge (\forall j \in \mathbb{N})(\forall k \in \mathbb{N})\Big(jk = p \Longrightarrow (j = 1 \vee k = 1)\Big) \tag{10.148}$$

is a sentence. It contains three variables, namely, p, j, and k. The variable p is free, but j and k are bound. So it should be possible to find an equivalent sentence that does not contain j and k at all. And, indeed, here is the sentence: "p is a prime number". This is not a proposition: its truth value depends on p. The sentence is true for p if p is a prime number, and is false if p is not a prime number.

6. The sentence

$$(\forall p \in \mathbb{Z})\left(p > 1 \wedge (\forall j \in \mathbb{N})(\forall k \in \mathbb{N})\Big(jk = p \Longrightarrow (j = 1 \vee k = 1)\Big)\right) \quad (10.149)$$

contains three variables, namely, p, j, and k. They are all bound. So the sentence is a proposition, and has a definite truth value. The sentence says "every integer is prime", which is of course false.

Important remark. Many students, when asked to define "prime number", answer by writing[57] (10.149). This, of course, is wrong. The students want to say "p is prime if and only if (10.148) holds", but instead they end up saying "every integer is prime", without understanding the difference. ***Please do not do that in your exam.*** □

Example 36. In each of the following sentences, the variable n is bound:

- "$(\forall n \in \mathbb{Z})n^2 - n$ is even". (In this case, the sentence itself says: "give n all possible integer values; for each such value, compute $n^2 - n$, see if it is even, and if the answer is "yes" for all value of n, then say "true"; otherwise say "false".)

- "$(\exists n \in \mathbb{Z})n^2 = 9$". (In this case, the sentence itself says: "give n all possible integer values; for each such value, compute n^2, and see if it is equal to 9; and if the answer is "yes" for at least one value of n, then say "true"; otherwise say "false".)

- "$\sum_{n=1}^{m} n^3 = \left(\frac{m(m+1)}{2}\right)^2$". (In this case, the sentence says: give n all possible integer values between 1 and m; for each such value, compute n^3; then add all the results, and see it the sum is equal to $\left(\frac{m(m+1)}{2}\right)^2$; if it is, say "true"; otherwise say "false". NOTE: In order to execute these instructions, you need to know who m is. So m is ***not*** a bound variable; the sentence does not generate a value for m. We have to tell the sentence who m is. So m is a ***free*** variable.)

- "$(\forall m \in \mathbb{N})\sum_{n=1}^{m} n^3 = \left(\frac{m(m+1)}{2}\right)^2$". (In this case, the sentence says: give m all possible natural number values; for each such value, do what was described in the previous example, to decide if "$\sum_{n=1}^{m} n^3 =$

[57] If you don't believe me, I can show you exams in which several students wrote exactly that. I don't understand why this happens, but it does.

$\left(\frac{m(m+1)}{2}\right)^2$" is true or not. If if the answer is "true" for all values of m, then say "true"; otherwise say "false".) So in this sentence m is also a bound variable.

Example 37. The same letter variable can occur in a sentence as both bound and free. So we really should not talk about a ***variable*** being free or being bound in a sentence: we should tall about an ***occurrence*** of a variable being free or bound.

For example, let S be the sentence

$$(\forall n \in \mathbb{Z})2|n \wedge n = 7$$

This very weird sentence says "every integer is even, and n is equal to 7." The letter n occurs three times in it, so ***there are three occurrences of*** n ***in*** S***.*** The first two are bound, and the third one is free. □

10.1.12 A convention about naming sentences: the expression $P(x)$

Sentences, like anything else, can be given letter names. And for sentences we will usually use capital letters. So, for example, a sentence could be called A, or B, or P, or Q, or X. But it will be convenient to sometimes use more complicated names, such as $P(x)$, or $P(x, y)$.

And we will adopt the following very useful convention:

We are allowed[a] to call a sentence $P(x)$, if x is free (i.e., not bound) in the sentence, that is, if the sentence does not contain an x-quantifier or any other expression (such as $\sum_{x=1} N$, or $\prod_{x=1}^{N}$) that assigns values to x.
Similarly, we are allowed to call a sentence $P(x, y)$, if the variables x and y are free (i.e., not bound) in the sentence, that is, if the sentence does not contain an x-quantifier or a y-quantifier or any other expression (such as $\sum_{x=1}^{N}$, or $\sum_{y=1}^{N}$, or $\prod_{x=1}^{N}$, $\prod_{y=1}^{N}$), that assigns values to x or y.

[a] I am saying "we are allowed to" rather "we have to". ***If a sentence has*** x ***as an open variable, we don't have to call it*** $P(x)$. We can call it P, if we want to.

Example 38.

- The following sentences can be called $P(x)$:

1. $2 + 2 = 4$,
2. $x > 5$,
3. $(x + 2)^2 = 7x - 3$,
4. $(y - x)^2 \geq 0$,
5. $y + 3 + y^2$,
6. $(\forall y \in \mathbb{R})(y - x)^2 \geq 0$,
7. $(\forall y \in \mathbb{R})y^2 \geq 0$,
8. $(x + y)^2 = x^2 + 2xy + y^2$,
9. $(\exists y \in \mathbb{R})x + y = 0$,
10. $(\forall y \in \mathbb{R})(\exists z \in \mathbb{R})y + z = x$.

- and the following sentences cannot be called $P(x)$:

1. $(\forall x \in \mathbb{R})(x + 1)^2 = x^2 + 2x + 1$,
2. $\sum_{x=1}^{5} x^2 = 55$,
3. $(\forall x \in \mathbb{R})x.0 = 0$.

- The following sentences can be called $P(x, y)$:

1. $2 + 2 = 4$,
2. $x > 5$,
3. $(x + 2)^2 = 7x - 3$,
4. $(y - x)^2 \geq 0$,
5. $y + 3 + y^2$,
6. $(x + y)^2 = x^2 + 2xy + y^2$.

- and the following sentences cannot be called $P(x, y)$:
- $(\forall y \in \mathbb{R})(y - x)^2 \geq 0$,
- $(\forall y \in \mathbb{R})y^2 \geq 0$,
- $(\forall y \in \mathbb{R})x + y = 0$,
- $(\forall x \in \mathbb{R})(x + 1)^2 = x^2 + 2x + 1$,
- $\sum_{x=1}^{5} x^2 = 55$,
- $(\exists y \in \mathbb{R})x + y = 0$,

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

(NOTE: If it bothers you that we are allowing using the name $P(x)$ for "$2 + 2 = 4$" and "$y + 3 + y^2$", even though these sentences have no x in them, the reason is very simple: all that matters is that x is not bound in these sentences. Whether it appears in them or not is irrelevant.)

10.1.13 Some problems

Problem 45. ***Prove*** each of the following propositions:

1. $(\forall n \in \mathbb{N})\left(\sum_{k=1}^{n} k = \frac{n(n+1)}{2} \Longrightarrow \sum_{k=1}^{n+1} k = \frac{(n+1)(n+2)}{2}\right)$.
2. $(\forall n \in \mathbb{N})\left(\sum_{k=1}^{n} k^2 = \frac{(2n+1)n(n+1)}{6} \Longrightarrow \sum_{k=1}^{n+1} k^2 = \frac{(2n+3)(n+1)(n+2)}{6}\right)$.
3. $(\forall n \in \mathbb{N})\left(\sum_{k=1}^{n} k^3 = \left(\frac{n(n+1)}{2}\right)^2 \Longrightarrow \sum_{k=1}^{n+1} k^3 = \left(\frac{(n+1)(n+2)}{2}\right)^2\right)$.
4. $(\forall n \in \mathbb{N})(n < 2^n \Longrightarrow n+1 < 2^{n+1})$.
5. $(\forall n \in \mathbb{N})\left(n^2 < 2^n + 2 \Longrightarrow (n+1)^2 < 2^{n+1} + 2\right)$.
6. $(\forall n \in \mathbb{N})\left(n^3 < 2^n + 257 \Longrightarrow (n+1)^3 < 2^{n+1} + 257\right)$.

10.2 Substitution

Substitution

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

  Example. If $P(x)$ is the sentence "$2 + 2 = 2 + x$", and t is the term "$1 + 1$", then $P(t)$ is the sentence "$2 + 2 = 2 + (1 + 1)$".

- ***We only allow the substitution of*** t ***for*** x ***in*** $P(x)$ ***when*** t ***is is free in*** $P(x)$***, in the sense that*** $P(x)$ ***does not contain a quantifier or any other expression that assigns values to any of the variables occurring in*** t.

Example. If $P(x)$ is the sentence "$(\exists y \in \mathbb{R})y = x$" (which is a sentence that contains x as an open variable, so we are allowed to call this sentence "$P(x)$"), and t is the term "y", then we are ***not*** allowed to substitute t for x in $P(x)$ and call the resulting sentence $P(y)$. □

In the following example, I will show you why the restriction on term substitution that we have just imposed in the box on "Substitution" is necessary.

Example 39. One of the rules of logic is Rule $\forall_{use}$, which says that

($\forall_{use}$) *If we have proved* $(\forall x \in S)P(x)$*, and* t *is a term, then we can go to* $P(t)$.

Suppose we allowed this for any sentence $P(x)$ and any term t, with no restictions. Then we would be able to take $P(x)$ to be the sentence "$(\exists y \in \mathbb{R})y = x$" and t to be the term "$y + 1$".

The sentence "$(\forall x \in \mathbb{R})P(x)$" says

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

It is easy to see that this sentence is in fact a true proposition. And it is easy to prove it. (Proof: Let x be an arbitrary real number. Pick y to be x. Then y is a witness for $(\exists y \in \mathbb{R})y = x$. So by Rule $\exists_{prove}$ we have proved $(\exists y \in \mathbb{R})y = x$ for arbirary $x \in \mathbb{R}$. Hence by Rule $\forall_{prove}$ we have proved that $(\forall x \in \mathbb{R})(\exists y \in \mathbb{R})y = x$.)

Now that we have proved "$(\forall x \in \mathbb{R})P(x)$", if we had no restriction on substitutions, we would be able to substitute the term t for x in $P(x)$, thus getting the sentence $P(t)$, that is, "$(\exists y \in \mathbb{R})y = y+1$". But this sentence is clearly false. So we do not want to be able to prove it from a true sentence. The only way to solve this problem is to avoid this kind of substitution. □

This is why, in order to avoid the problem that we presented in Example 39, we impose the restriction explained earlier: ***in a sentence*** $P(x)$ ***we can only substitute for*** x ***a term*** t ***that does not contain any variables that are bound in*** $P(x)$.

So, for example, we can conclude from the sentence "$(\forall x \in \mathbb{R})(\exists y \in \mathbb{R})y \neq x$" that "$(\exists y \in \mathbb{R})y \neq 5$", or "$(\exists y \in \mathbb{R})y \neq x$", or "$(\exists y \in \mathbb{R})y \neq x^2 + z + 35$", but not "$(\exists y \in \mathbb{R})y \neq y$".

10.3 Forming sentences: the grammar of formal language

The English language has a ***grammar***, i.e., a set of rules that restrict what combinations of words are acceptable ("grammatically correct") sentences. For example, "cows eats grass" is not a grammatically correct English sentence, because the subject is the noun "cows", which is plural, and the verb is "eats", which is singular[58]

English grammar is very complicated, with lots of rules, an enormous number of exceptions, and many cases where it is unclear whether something is a grammatically correct sentence. (For example, people argue about

[58] English grammar is crazy! Most nouns form their plural by adding an "s" at end, so the plurals of "cow", "duck", "table" are "cows", "ducks", "tables". But for most verbs it's the other way around: the singular ends with an "s" (as in "eats", "swims", "walks") and the plural is without the "s" (as in "eat", "swim", "walk"), so "cows eat grass" and "ducks swim" are grammatically correct, but "cows eats grass" and "ducks swims" are not. Go figure!

whether a sentence such as "He is determined to completely destroy the evidence", containing a split infinitive, is correct or not.)

Formal mathematical language has a very simple grammar. Here is the part of formal language grammar that has to do with the formation of sentences. (There is another part that deals with the formation of atomic sentences. That will be discussed later.)

- Sentences are formed by combining atomic sentences, connectives, and parentheses.

- Atomic sentences are sentences.
- If A, B are sentences, then we can form the sentences $A \wedge B$, $A \vee B$, $A \Longrightarrow B$, and $A \Longleftrightarrow B$.
- If A is a sentence, then we can form the sentence $\sim A$, the ***negation*** of A,
- If A is a sentence and Q is a quantifier, then we can form the sentence QA, known as an *existential quantification of* A, if Q is an existential quantifier, and as a *universal quantification of* A, if Q is a universal quantifier[59]

- When a sentence A is combined with other sentences or connectives to form another sentence, then: if A is of the form $P \Longrightarrow Q$, or $P \wedge Q$, or $P \vee Q$, or $P \Longleftrightarrow Q$, then A has to be enclosed in parentheses before we form the combination.

Example 40. Let us say that "if n is an integer then if n is even then $n+1$ is odd". To say this, we use the atomic sentences "$2|n$" ("n is even") and "$2|n+1$" ("$n+1$ is even") and the connectives "$\sim$" and "$(\forall n \in \mathbb{Z})$". We negate "$2|n+1$" to form the sentence "$\sim 2|n+1$", which says "$n+1$ is odd". Then we combine "$2|n$" and "$\sim 2|n+1$" using the connective "$\Longrightarrow$", and form the sentence "$2|n \Longrightarrow \sim 2|n+1$" ("if n is even then $n+1$ is odd"). Finally, in order to assert that "$2|n \Longrightarrow \sim 2|n+1$" is true for every integer n, we quantify "$2|n \Longrightarrow \sim 2|n+1$" by writing the quantifier "$(\forall n \in \mathbb{Z})$" to its left. But before we do that, since the sentence "$2|n \Longrightarrow \sim 2|n+1$" is of the form $A \Longrightarrow B$, we enclosed it in parentheses, by writing "$(2|n \Longrightarrow \sim 2|n+1)$" . The final result is the sentence

$$(\forall n \in \mathbb{Z})(2|n \Longrightarrow \sim 2|n+1)\,.$$

[59] ***Quantifers*** were discussed in Section 10.4.2, on page 161. Recall that: the symbols "$\forall$" and "$\exists$" are the ***quantifier symbols***. Using these symbols, we can form expressions such as "$(\forall x)$" and "$(\exists x)$", called ***unrestricted quantifiers***, and "$(\forall x \in S)$" and "$(\exists x \in S)$", called ***restricted quantifiers***.

This sentence says precisely what we want, i.e., that the statement "if n is even then $n+1$ is odd" is true for every integer n.

Example 41. Suppose we want to say

> if a natural number p has the property that whenever two integers a, b are such that p divides ab it follows that p divides a or p divides b, then p is a prime number or $p = 1$

in formal language.

We observe first that sentence clearly says that something is true for every natural number p, so the sentence is of the form "$(\forall n \in \mathbb{N})A$". Now, A is of the form $B \Longrightarrow C$, where B is the sentence "p has the property that whenever two integers a, b are such that p divides ab it follows that p divides a or p divides b", and C is the sentence "p is a prime number or $p = 1$".

Then, if formal language, A says

$$(\forall a \in \mathbb{Z})(\forall b \in \mathbb{Z})\Big(p|ab \Longrightarrow (p|a \vee p|b)\Big)\,,$$

and B says

$$p = \text{is a prime number } \vee p = 1\,.$$

So our sentence says

$$(\forall p \in \mathbb{N})\Big((\forall a \in \mathbb{Z})(\forall b \in \mathbb{Z})\Big(p|ab \Longrightarrow (p|a \vee p|b)\Big) \Longrightarrow (p \text{ is a prime number} \vee p = 1)\Big).$$

This is not yet a completely formal sentence, because it has the words "is a prime number". In order to get a completely formal sentence, we can substitute for "p is a prime number" the meaning of "p is a prime number" in formal language, that is,

$$(\forall k \in \mathbb{N})\Big(k|p \Longrightarrow (k = 1 \vee k = p)\Big)\,.$$

The result is

$$(\forall p \in \mathbb{N})\Bigg((\forall a \in \mathbb{Z})(\forall b \in \mathbb{Z})\Big(p|ab \Longrightarrow (p|a \vee p|b)\Big)$$

$$\Longrightarrow \Big((\forall k \in \mathbb{N})\big(k|p \Longrightarrow (k = 1 \vee k = p)\big) \vee p = 1\Big)\Bigg)\,.$$

Notice that this sentence contains several letter variables, namely, p, a, b, and k, but they are all bound variables, so the sentence is a proposition. And we can see this by rephrasing the sentence without using any letter variables at all. as follows:

if a natural number has the property that whenever it divides the product of two integers it follows that it divides one of them, then the natural number is either one or a prime number.

10.4 How sentences are constucted

10.4.1 The seven logical connective symbols

There are ***seven*** logical connectives[a]. They are:

1. $\sim$: negation ("it's not the case that"),
2. $\wedge$: conjunction ("and"),
3. $\vee$: disjunction ("or"),
4. $\Longrightarrow$: implication ("implies", "if $\cdots$ then"),
5. $\Longleftrightarrow$: biconditionsl ("if and only if"),
6. universal quantifiers,
7. existential quantifiers.

[a] A Logical connective is a symbol that is used to combine sentences to form new sentences.

10.4.2 The quantifiers

- A quantifier is an expression

$$(\forall x) \quad \text{or} \quad (\forall x \in S) \quad \text{or} \quad (\exists x) \quad \text{or} \quad (\exists x \in S)\,.$$

  where x is a variable and S is a set.
- The symbol "$\forall$" is called the universal quantifier symbol.
- The symbol "$\exists$" is called the existential quantifier symbol.
- "$(\forall x)$" is an unrestricted universal quantifier,
- "$(\forall x \in S)$" is a restricted universal quantifier.
- "$(\exists x)$" is an unrestricted existential quantifier,
- "$(\exists x \in S)$" is a restricted existential quantifier,

10.4.3 Sentence types

Every mathematical sentence is of one, and only one, of the following eight types:

1. atomic,
2. negation ("$\sim A$"),
3. conjunction ("$A \wedge B$"),
4. disjunction ("$A \vee B$"),
5. implication ("$A \Longrightarrow B$"),
6. biconditional ("$A \Longleftrightarrow B$"),
7. universal ("$(\forall x)P(x)$" or "$(\forall x \in S)P(x)$")
8. existential ("$(\exists x)P(x)$" or "$(\exists x \in S)P(x)$").

10.5 Forming sentences

- Sentences are formed by combining atomic sentences, connectives, and parentheses.
- Atomic sentences are sentences.
- If A, B are sentences, then we can form the sentences $A \wedge B$, $A \vee B$, $A \Longrightarrow B$, and $A \Longleftrightarrow B$.
- If A is a sentence, then we can form the sentence $\sim A$, the ***negation*** of A,
- If A is a sentence and Q is a quantifier, then we can form the sentence QA, known as an *existential quantification of* A, if Q is an existential quantifier, and as a *universal quantification of* A, if Q is a universal quantifier.
- When a sentence A is combined with other sentences or connectives to form another sentence, then: if A is of the form $P \Longrightarrow Q$, or $P \wedge Q$, or $P \vee Q$, or $P \Longleftrightarrow Q$, then A has to be enclosed in parentheses before we form the combination.

Example 42. Let us say that "if n is an integer then if n is even then $n+1$ is odd". To say this, we use the atomic sentences "$2|n$" ("n is even") and "$2|n+1$" ("$n+1$ is even") and the connectives "$\sim$" and "$(\forall n \in \mathbb{Z})$". We negate "$2|n+1$" to form the sentence "$\sim 2|n+1$", which says "$n+1$ is odd". Then we combine "$2|n$" and "$\sim 2|n+1$" using the connective "$\Longrightarrow$", and form the sentence "$2|n \Longrightarrow \sim 2|n+1$" ("if n is even then $n+1$ is odd"). Finally, in order to assert that "$2|n \Longrightarrow \sim 2|n+1$" is true for every integer n, we quantify "$2|n \Longrightarrow \sim 2|n+1$" by writing the quantifier "$(\forall n \in \mathbb{Z})$" to its left. But before we do that, since the sentence "$2|n \Longrightarrow \sim 2|n+1$" is of the form $A \Longrightarrow B$, we enclosed it in parentheses, by writing "$(2|n \Longrightarrow \sim 2|n+1)$" . The final result is the sentence

$$(\forall n \in \mathbb{Z})(2|n \Longrightarrow \sim 2|n+1)\,.$$

This sentence says precisely what we want, i.e., that the statement "if n is even then $n+1$ is odd" is true for every integer n.

10.5.1 When do we put parentheses?

When a sentence A is combined with other sentences or connectives to form another sentence, then: if A is of the form $P \Longrightarrow Q$, or $P \wedge Q$, or $P \vee Q$, or $P \Longleftrightarrow Q$, then A has to be enclosed in parentheses before we form the combination.

Example 43. Suppose you want to say that

If an integer n is even, then n^2 is divisible by 4.

You start with the atomic sentences "$2|n$" ("n is even") and "$4|n^2$" ("n^2 is divisible by 4").

Then you combine these sentences using the implication connective, and get

$$2|n \Longrightarrow 4|n^2\,, \tag{10.150}$$

that is, "if n is even, then n^2 is divisible by 4".

Finally, it is clear that the sentence is intended to be an assertion for every integer n, so you quantify it with a universal quantifier. But before you quantify, you have to remember that (10.150) is an implication, that is, a sentence of the form $A \Longrightarrow B$. So ***before you quantify it, you have to enclose it in parentheses***.

The final result is the proposition

$$(\forall n \in \mathbb{Z})(2|n \Longrightarrow 4|n^2)\,. \tag{10.151}$$

What would have happened if you had not put the parentheses? You would have ended up with

$$(\forall n \in \mathbb{Z})2|n \Longrightarrow 4|n^2\,, \tag{10.152}$$

which is a completely different sentence! Sentence (10.152) says that the sentence "$(\forall n \in \mathbb{Z})2|n$" implies the sentence "$4|n^2$". In other words, (10.152) says: "if every integer is even, then n is divisible by 4". This is not even a proposition, because the third "n" is an open variable. □

10.6 The 14 logical rules

Here is the list of the fourteen logical rules.

1

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.

2

Rule for proving a conjunction (Rule $\wedge_{prove}$)
If P, Q are sentences, and you have proved P and have proved Q, then you are allowed to go to $P \wedge Q$,

3

Rule for using an implication (Rule $\Longrightarrow_{use}$, a.k.a. Modus Ponens)
If P, Q are sentences, and you have proved $P \Longrightarrow Q$ and have proved P, then you are allowed to go to Q.

4

Rule for proving an implication (Rule $\Longrightarrow_{prove}$)
If P, Q are sentences, and you have proved Q assuming P, then you are allowed to go to $P \Longrightarrow Q$.

5

Rule for using a biconditional (Rule $\Longleftrightarrow_{use}$)
If P, Q are sentences, and you have proved $P \Longleftrightarrow Q$, then you can go to $P \Longrightarrow Q$ and to $Q \Longrightarrow P$.

6

Rule for proving a biconditional (Rule $\Longleftrightarrow_{prove}$)

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

7

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

If P, Q, R are sentences, and you have proved $P \vee Q$, $P \Longrightarrow R$, and $Q \Longrightarrow R$, then you can go to R.

8

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

Suppose P and Q are sentences. Then, if you have proved $\sim P \Longrightarrow Q$ or $\sim Q \Longrightarrow P$ then you can go to $P \vee Q$.

9

The proof by contradiction rule

(I) If, assuming A, we prove C, which is a contradiction, then we can go to $\sim A$.

(II) If, assuming $\sim A$, we prove C, which is a contradiction, then we can go to A.

10

Rule for using universal sentences (Rule $\forall_{use}$, a.k.a. the "universal specialization rule")

If $P(x)$ is a sentence and t is a term that does not contain any variables that are bound in $P(x)$, then

- if you have proved $(\forall x)P(x)$, you can go to $P(t)$;
- If you have proved that $(\forall x \in S)P(x)$, and that $t \in S$, then you can go to $P(t)$.

11

Rule for proving universal sentences (Rule $\forall_{prove}$, a.k.a. the "universal generalization rule")

(I) If, starting with "Let x be arbitrary", you prove $P(x)$, then you are allowed to go to $(\forall x)P(x)$.

(II) If, starting with "Let $x \in S$ be arbitrary", or "Let x be an arbitrary member of S", you prove $P(x)$, then you are allowed to go to $(\forall x \in S)P(x)$.

12

Rule for using existential sentences (Rule $\exists_{use}$, a.k.a. the "existential specialization rule")

If $P(x)$ is a sentence, and the letter a is not in use as the name of anything, then:

- If you have proved $(\exists x)P(x)$, then you can introduce a witness for $(\exists x)P(x)$, and call it a, so that this new object will satisfy $P(a)$.
- In addition, if S is a set, and you have proved that $(\exists x \in S)P(x)$, then you can stipulate that $a \in S$ as well.

13

Rule for proving existential sentences, (Rule $\exists_{prove}$, a.k.a. the "existential generalization rule")

Suppose $P(x)$ is a sentence and w is a term that does not contains any variables that are bound in $P(x)$. Then

(I) If w is a witness for $(\exists x)P(x)$ (i.e., if $P(w)$), then you can go to $(\exists x)P(x)$.

(II) If w is a witness for $(\exists x \in S)P(x)$ (i.e., if $w \in S$ and $P(w)$), then you can go to $(\exists x \in S)P(x)$.

14

"Substitution of equals for equals" (Rule SEE)

If $P(x)$ is a sentence, s and t are terms, and you have proved $s = t$ or $t = s$, and you have also proved $P(s)$, then you can go to $P(t)$.

10.7 Using the rules: examples of "pure logic" proofs

In this section I will illustrate how to use the rules of logic by giving several examples of proofs that involve ***no mathematics, only the rules of logic***.

Example 44. Let A, B, C be propositions. ***Prove***, using the rules of logic, that

$$\big(A \Longrightarrow (B \Longrightarrow C)\big) \Longleftrightarrow \Big((A \wedge B) \Longrightarrow C\Big)\,. \tag{10.153}$$

Solution.

We want to prove a biconditional sentence. For that purpose. we use Rule $\Longleftrightarrow_{prove}$*: to prove* $P \Longleftrightarrow Q$*, prove* $P \Longrightarrow Q$ *and* $Q \Longrightarrow P$.

Proof of "$\big(A \Longrightarrow (B \Longrightarrow C)\big) \Longrightarrow \Big((A \wedge B) \Longrightarrow C\Big)$".

Assume $\boxed{A \Longrightarrow (B \Longrightarrow C)}$. *We want to prove* $(A \wedge B) \Longrightarrow C$.

Assume $\boxed{A \wedge B}$. *We want to prove* C.
It follows from $A \wedge B$ that A. [Rule $\wedge_{use}$]
It follows from A and $A \Longrightarrow (B \Longrightarrow C)$ that $B \Longrightarrow C$. [Rule $\Longrightarrow_{use}$]
It follows from $A \wedge B$ that B, [Rule $\wedge_{use}$]
It follows from B and $B \Longrightarrow C$ that $\boxed{C}$. [Rule $\Longrightarrow_{use}$]

Since we have proved C assuming $A \wedge B$, we conclude that $\boxed{(A \wedge B) \Longrightarrow C}$. [Rule $\Longrightarrow_{prove}$]

Since we have proved $(A \wedge B) \Longrightarrow C$ assuming $A \Longrightarrow (B \Longrightarrow C)$, we can go to

$$\Big(A \Longrightarrow (B \Longrightarrow C)\Big) \Longrightarrow \big((A \wedge B) \Longrightarrow C\big), \tag{10.154}$$

completing the proof of "$\Big(A \Longrightarrow (B \Longrightarrow C)\Big) \Longrightarrow \big((A \wedge B) \Longrightarrow C\big)$".

Proof of "$\Big((A \wedge B) \Longrightarrow C\Big) \Longrightarrow \Big(A \Longrightarrow (B \Longrightarrow C)\Big)$".

Assume $\boxed{(A \wedge B) \Longrightarrow C}$. *We want to prove* $A \Longrightarrow (B \Longrightarrow C)$.

Assume $\boxed{A}$. *We want to prove* $B \Longrightarrow C$.

Assume $\boxed{B}$. *We want to prove* C.
Since we have A and B, we can go to $A \wedge B$. [Rule $\wedge_{prove}$]
Since we have $A \wedge B$ and $(A \wedge B) \Longrightarrow C$, we can go to $\boxed{C}$. [Rule $\Longrightarrow_{use}$]

Since we have proved C assuming B, we can go to $\boxed{B \Longrightarrow C}$. [Rule $\Longrightarrow_{use}$]

Since we have proved $B \Longrightarrow C$ assuming A, we can go to $\boxed{A \Longrightarrow (B \Longrightarrow C)}$. [Rule $\Longrightarrow_{use}$]

Since we have proved $A \Longrightarrow (B \Longrightarrow C)$ assuming $(A \wedge B) \Longrightarrow C$, we can go to

$$\big((A \wedge B) \Longrightarrow C\big) \Longrightarrow \Big(A \Longrightarrow (B \Longrightarrow C)\Big), \tag{10.155}$$

completing the proof of "$\big((A \wedge B) \Longrightarrow C\big) \Longrightarrow \Big(A \Longrightarrow (B \Longrightarrow C)\Big)$".

Since we have proved both implications (10.154) and (10.155), we can conclude from Rule $\Longleftrightarrow_{prove}$ that

$$\Big(A \Longrightarrow (B \Longrightarrow C)\Big) \Longleftrightarrow \big((A \wedge B) \Longrightarrow C\big)\,, \tag{10.156}$$

Q.E.D.

Example 45. Let A be a proposition. ***Prove***, using the rules of logic, that

$$A \Longleftrightarrow \sim\sim A\,. \tag{10.157}$$

Remark 11. Formula (10.157) captures our intuition that $\sim\sim A$ is "the same as A". □

Solution.

In order to prove the biconditional sentence $A \Longleftrightarrow \sim\sim A$, we will prove the implications $A \Longrightarrow \sim\sim A$ and $\sim\sim A \Longrightarrow A$.

Proof of "$A \Longrightarrow \sim\sim A$".
Assume A. *We want to prove* $\sim\sim A$.
We will prove $\sim\sim A$ *by contradiction.*
Assume $\sim A$.
Then $A \wedge \sim A$. [Rule $\wedge_{prove}$]
Since we have proved a contradiction assuming $\sim A$, we get $\sim\sim A$. 11100000 [Proof by contradiction rule]
Since we have proved $\sim\sim A$ assuming A, we get $\boxed{A \Longrightarrow \sim\sim A}$. [Rule $\Longrightarrow_{prove}$]
Now assume $\sim\sim A$. *We want to prove* A.
We will prove A *by contradiction.*
Assume $\sim A$.
Then $\sim A \wedge \sim\sim A$. [Rule $\wedge_{prove}$]
Since we have proved a contradiction assuming $\sim A$, we get A.
[Proof by contradiction rule]
Since we have proved A assuming $\sim\sim A$, we get $\boxed{\sim\sim A \Longrightarrow A}$. [Rule $\Longrightarrow_{prove}$]
Since we have proved $A \Longrightarrow \sim\sim A$ and $\sim\sim A \Longrightarrow A$, we get $A \Longleftrightarrow \sim\sim A$.

Q.E.D.

Example 46. Let A, B be propositions. ***Prove***, using the rules of logic, that

$$A \Longrightarrow (A \vee B)\,. \tag{10.158}$$

Remark 12. Formula (10.158) captures the obvious idea expressed that if "A" is true then "$A \vee B$" must be true. □

Solution.

In order to prove the implication "$A \Longrightarrow (A \vee B)$", we will use Rule $\Longrightarrow_{prove}$.

For that purpose, we will assume that A and prove that $A \vee B$.
Assume that A. *We want to prove that $A \vee B$.*
In order to prove "$A \vee B$" we will use Rule $\vee_{prove}$.
For that purpose, we will assume that $\sim A$ and prove that B.
Assume that $\sim A$. *We want to prove that B.*
We will prove that B by contradiction.
Assume that $\sim B$.
$A \wedge \sim A$.
Since we proved a contradiction assuming that $\sim B$, we conclude that B.
[Proof by contradiction rule]
Since we have proved that B assuming that $\sim A$, we conclude that $A \vee B$.[$\vee_{prove}$]
Since we have proved that $A \vee B$ assuming that A, we conclude that $\boxed{A \Longrightarrow (A \vee B)}$.
[Rule $\Longrightarrow_{prove}$]
Q.E.D.

Example 47. Let A, B be propositions. ***Prove***, using the rules of logic, that

$$(A \Longrightarrow B) \Longleftrightarrow (\sim A \vee B)\,. \tag{10.159}$$

Remark 13. Formula (10.159) captures the idea expressed by the truth table for the implication: "$A \Longrightarrow B$" is true when A is false or B is true, which means that "$A \Longrightarrow B$" is "the same as '$\sim A \vee B$' ". □

Solution.

In order to prove the biconditional sentence "$(A \Longrightarrow B) \Longleftrightarrow (\sim A \vee B)$", we will prove the implications "$(A \Longrightarrow B) \Longrightarrow (\sim A \vee B)$" and "$(\sim A \vee B) \Longrightarrow (A \Longrightarrow B)$".

Proof of "$(A \Longrightarrow B) \Longrightarrow (\sim A \vee B)$".
Assume "$A \Longrightarrow B$". *We want to prove "$\sim A \vee B$".*
For this purpose, we will use Rule $\vee_{prove}$. We will prove "$\sim B \Longrightarrow \sim A$".
Assume that $\sim B$. *We want to prove that $\sim A$.*
We will prove that $\sim A$ by contradiction.
Assume that A.
B [Rule $\Longrightarrow_{use}$]
$B \wedge \sim B$ [Rule $\wedge_{prove}$]
Since we have proved a contradiction assuming that A, we get $\boxed{\sim A}$. [PCR]
Since we have proved that $\sim A$ assuming that $\sim B$, we get $\boxed{\sim A \vee B}$. [$\vee_{prove}$]
Since we proved "$\sim A \vee B$" assuming "$A \Longrightarrow B$", we get $\boxed{(A \Longrightarrow B) \Longrightarrow (\sim A \vee B)}$.
[Rule $\Longrightarrow_{prove}$]

Proof of "$(\sim A \vee B) \Longrightarrow (A \Longrightarrow B)$".
Assume "$\sim A \vee B$". *We want to prove "$A \Longrightarrow B$".*

Assume that A. *We want to prove that* B.
To prove that B *using "*$\sim A \vee B$*" we use Rule* $\vee_{use}$.
For that purpose, we must prove that $\sim A \Longrightarrow B$ *and* $B \Longrightarrow B$.
Proof that $\sim A \Longrightarrow B$.
Assume that $\sim A$. *We want to prove that* B.
Assume that $\sim B$.
$A \wedge \sim A$. [Rule $\wedge_{prove}$]
Since we have proved a contradiction assuming that $\sim B$, we get B. [PCR]
Since we have proved "B" assuming "$\sim A$". we get $\boxed{\sim A \Longrightarrow B}$. [$\Longrightarrow_{prove}$]
Proof that $B \Longrightarrow B$.
Assume that B.
Then B.
So $\boxed{B \Longrightarrow B}$. [Rule $\Longrightarrow_{prove}$]
Since we know that $\sim A \vee B$, and we have proved that $\sim A \Longrightarrow B$ and that $B \Longrightarrow B$, we can conclude that B. [Rule $\vee_{use}$]
Since we have proved "B" assuming "A", we conclude that $A \Longrightarrow B$. [Rule $\Longrightarrow_{prove}$]
Since we have proved that $A \Longrightarrow B$ assuming "$\sim A \vee B$", we conclude that $\boxed{(\sim A \vee B) \Longrightarrow (A \Longrightarrow B)}$. [Rule $\Longrightarrow_{prove}$]
Since we have proved that $(\sim A \vee B) \Longrightarrow (A \Longrightarrow B)$ and $(A \Longrightarrow B) \Longrightarrow (\sim A \vee B)$, we conclude that $\boxed{(A \Longrightarrow B) \Longrightarrow (\sim A \vee B)}$. [Rule $\Longleftrightarrow_{prove}$]
Q.E.D.

Example 48. Let A, B be propositions. ***Prove*** that the proposition

$$A \Longrightarrow (A \wedge B) \tag{10.160}$$

cannot be proved using only the rules of logic.

Solution. If it was possible to prove (10.160) using only the rules of logic, then it would follow that no matter which propositions we substitute for the letters A, B, the resulting proposition is true.

However, suppose we take A to be the sentence "1 is odd", and B to be the sentence "1 is even"'. Then B is false, so "$A \wedge B$" is false. But A is true, so "$A \Longrightarrow (A \wedge B)$" is false. Therefore "$A \Longrightarrow (A \wedge B)$" cannot be proved using only the rules of logic. □

Example 49. Let $P(x)$, $Q(x)$, be one-variable predicates, and let S be a set.

1. ***Prove***, using the rules of logic, the sentence

$$(\forall x \in S)(P(x) \wedge Q(x)) \Longleftrightarrow \Big((\forall x \in S)P(x) \wedge (\forall x \in S)Q(x)\Big) \tag{10.161}$$

(Here is an example: suppose S is the set of all people, "$P(x)$" stands for "x likes tea" and "$Q(x)$" stands for "x likes coffee". Then the sentence "$(\forall x \in S)(P(x) \wedge Q(x))$" says "everybody likes tea and coffee", and the sentence "$(\forall x \in S)P(x) \wedge (\forall x \in S)Q(x)$" says "everybody likes tea and everybody likes coffee". It is clear that both sentences say the same thing, so it is obvious that they are equivalent, so that the sentence (10.161) is true.)

2. ***Prove*** that the sentence

$$(\forall x \in S)(P(x) \vee Q(x)) \iff \Big((\forall x \in S)P(x) \vee (\forall x \in S)Q(x)\Big) \quad (10.162)$$

cannot be proved using the rules of logic. (HINT: Find an example of a pair of predicates $P(x)$, $Q(x)$ for which (10.162) is false.)

Solution. First, we prove (10.161).

Sentence (10.161) is a biconditional, of the form $A \iff B$. So, in order to prove it, we will use Rule $\iff_{prove}$, and prove both $A \implies B$ and $B \implies A$.

Proof of "$(\forall x \in S)(P(x) \wedge Q(x)) \implies \Big((\forall x \in S)P(x) \wedge (\forall x \in S)Q(x)\Big)$".

Assume that

$$(\forall x \in S)(P(x) \wedge Q(x))\,. \quad (10.163)$$

We want to prove that $(\forall x \in S)P(x) \wedge (\forall x \in S)Q(x)$.

For this purpose, we will use Rule $\wedge_{prove}$, and prove the sentences $(\forall x \in S)P(x)$ and $(\forall x \in S)Q(x)$.

Proof of $(\forall x \in S)P(x)$:

Let u be an arbitrary member of S.

Then $P(u) \wedge Q(u)$. [Rule $\forall_{use}$, from (10.163)]

Therefore $P(u)$. [Rule $\wedge_{use}$, from $P(u) \wedge Q(u)$.

So we have proved $P(u)$ for an arbitary $u \in S$, and then $\boxed{(\forall x \in S)P(x)}$. [Rule $\forall_{prove}$]

Proof of $(\forall x \in S)Q(x)$:

Let u be an arbitrary member of S.

Then $P(u) \wedge Q(u)$. [Rule $\forall_{use}$, from (10.163)]

Therefore $Q(u)$. [Rule $\wedge_{use}$, from $P(u) \wedge Q(u)$.

So we have proved $Q(u)$ for an arbitary $u \in S$, and then $\boxed{(\forall x \in S)Q(x)}$. [Rule $\forall_{prove}$]

We have proved $(\forall x \in S)P(x)$ and $(\forall x \in S)Q(x)$. Therefore

$$(\forall x \in S)P(x) \wedge (\forall x \in S)P(x)\,, \tag{10.164}$$

by Rule $\wedge_{prove}$.

Since we have proved 10.164) assuming (10.163), we can go to

$$(\forall x \in S)(P(x) \wedge Q(x)) \Longrightarrow \Big((\forall x \in S)P(x) \wedge (\forall x \in S)Q(x)\Big)\,, \tag{10.165}$$

completing the proof of (10.165).

Proof of "$\Big((\forall x \in S)P(x) \wedge (\forall x \in S)Q(x)\Big) \Longrightarrow (\forall x \in S)(P(x) \wedge Q(x))$"

Assume

$$(\forall x \in S)P(x) \wedge (\forall x \in S)Q(x)\,. \tag{10.166}$$

We want to prove $(\forall x \in S)(P(x) \wedge Q(x))$.

Let u be an arbitrary member of S.

It follows from (10.166) by Rule $\wedge_{use}$ that $(\forall x \in S)P(x)$.

And it also follows that $(\forall x \in S)Q(x)$.

Since $(\forall x \in S)P(x)$, and $u \in S$, it follows by Rule $\forall_{use}$ that $P(u)$.

Since $(\forall x \in S)Q(x)$, and $u \in S$, it follows by Rule $\forall_{use}$ that $Q(u)$.

Since we have proved $P(u)$ and $Q(u)$, it follows by Rule $\wedge_{prove}$ that $\boxed{P(u) \wedge Q(u)}$.

Since we have proved $P(u) \wedge Q(u)$ for arbitrary u in S, it follows by Rule $\forall_{prove}$ that

$$(\forall x \in S)(P(x) \wedge Q(x))\,. \tag{10.167}$$

Since we have proved (10.167) assuming (10.166), it follows from Rule $\Longrightarrow_{prove}$ that

$$\Big((\forall x \in S)P(x) \wedge (\forall x \in S)Q(x)\Big) \Longrightarrow (\forall x \in S)(P(x) \wedge Q(x))\,. \tag{10.168}$$

completing the proof of (10.168).

Since we have proved (10.165) and (10.168), it follws by Rule $\Longleftrightarrow_{prove}$ that

$$(\forall x \in S)(P(x) \wedge Q(x)) \Longleftrightarrow \Big((\forall x \in S)P(x) \wedge (\forall x \in S)Q(x)\Big)\,, \quad (10.169)$$

Q.E.D.

We now prove that (10.162) cannot be proved. We do this by exhibitng examples of a set S and predicates $P(x)$, $Q(x)$ for which (10.162) is false.

Let S be the set of all people. Let "$P(x)$" stand for "x likes tea" and "$Q(x)$" stand for "x likes coffee". Then the sentence "$(\forall x \in S)(P(x) \vee Q(x))$" says "everybody likes tea or coffee", and the sentence "$(\forall x \in S)P(x) \vee (\forall x \in S)Q(x)$" says "everybody likes tea or everybody likes coffee". It is clear that both sentences say totally different things. The sentence "everybody likes tea or everybody likes coffee" is certainly false, because it is a disjunction "everybody likes tea $\vee$ everybody likes coffee", and both disjuncts ("everybody likes tea" and "everybody likes coffee") are false, so the disjunction is false.

10.8 Some problems, with solutions

Problem 46. For the sentence

$$(\forall n \in \mathbb{Z})(\exists m \in \mathbb{Z})m > n\,, \quad (10.170)$$

i. Translate the sentence into reasonable English.

ii. List all the variables that occur in the sentence, and indicate which ones are free (i.e. open) and which ones are bound (i.e., dummy, or closed). If a variable occcurs in the sentence more than once, it may happen that some of the occurrences are free and others are boiund. If this happens, say it.

iii. Indicate whether the sentence is a proposition (i.e., has no open variables) or not.

iv. If the sentence is a proposition, then

   a. ***indicate*** whether it is true or false,
   b. ***prove*** the assertion that you made to answer part a.

Solution.

i. Sentence (10.170) says: "for every integer n there exists an integer m such that $m > n$".

ii. The variables occurring in (10.170) are m and n. Both are bound. The sentence has no free variables.

iii. The sentence is a proposition.

iv.a The sentence is ***true***.

iv.b *Proof of (10.170):*

> Let $n \in \mathbb{Z}$ be arbitrary.
> *We want to prove "$(\exists m \in \mathbb{Z})m > n$", and for that purpose we are going to find a witness.*
> Chooe $w = n + 1$.
> *Why do I choose w this way? Because it works. How do I know it works? In this case, it is quite obvious: I need an integer greater than n, so $n + 1$ is a natural choice.*
> Then $w \in \mathbb{Z}$ and $w > n$.
> So w is a witness for "$(\exists m \in \mathbb{Z})m > n$".
> Hence $(\exists m \in \mathbb{Z})m > n$. [Rule $\exists_{prove}$]

So we have proved "$(\exists m \in \mathbb{Z})m > n$" for arbitrary $n \in \mathbb{Z}$.

Hence $\boxed{(\forall n \in \mathbb{Z})(\exists m \in \mathbb{Z})m > n}$. [Rule $\forall_{prove}$] **Q.E.D.**

Problem 47. For the sentence

$$(\exists m \in \mathbb{Z})m > n\,, \tag{10.171}$$

perform exactly the same tasks i., ii., iii., and iv. (a. and b.) as in Problem 46.

Solution.

i. Sentence (10.171) says: "there exists an integer m such that $m > n$".

ii. The variables occurring in (10.171) are m and n. The variable m is bound, and m is free.

iii. The sentence is a not a proposition.

iv.a,b Since the sentence is not a proposition, questions [iv.a] and [iv.b] do not apply.

Problem 48. For the sentence

$$(\exists m \in \mathbb{Z})(\forall n \in \mathbb{Z})m > n\,, \tag{10.172}$$

perform exactly the same tasks i., ii., iii., and iv. (a. and b.) as in Problem 46.

Solution.

i. Sentence (10.172) says: "there exists an integer m such that m is larger than every integer".
ii. The variables occurring in (10.172) are m and n. Both are bound. The sentence has no free variables.
iii. The sentence is a proposition.
iv.a The sentence is ***false***.
iv.b *Proof that (10.172) is false:*

We want to prove that $\sim (\exists m \in \mathbb{Z})(\forall n \in \mathbb{Z})m > n$.

We will do it by contradiction.

Assume that $(\exists m \in \mathbb{Z})(\forall n \in \mathbb{Z})m > n$.

Pick a witness w, so $w \in \mathbb{Z}$ and $(\forall n \in \mathbb{Z})m > n$. [Rule $\exists_{use}$]

Then $w > w + 1$. [Rule $\forall_{use}$].

But $(\forall n \in \mathbb{Z})n \leq n + 1$.

So $w \leq w + 1$. [Rule $\forall_{use}$]

Hence $\sim w > w + 1$.

So $w > w + 1 \wedge \sim w > w + 1$ [Rule $\wedge_{prove}$]

And "$w > w + 1 \wedge \sim w > w + 1$" is a contradiction.

So we have proved a contradiction assuming that $(\exists m \in \mathbb{Z})(\forall n \in \mathbb{Z})m > n$.

Therefore $\boxed{\sim (\exists m \in \mathbb{Z})(\forall n \in \mathbb{Z})m > n}$. [Proof by contradiction rule]
Q.E.D.

Problem 49. For the sentence

$$(\forall n \in \mathbb{Z})(2|n \Longrightarrow 4|n^2), \tag{10.173}$$

perform exactly the same tasks i., ii., iii., and iv. (a. and b.) as in Problem 46.

Solution.

i. Sentence (10.173) says: "the square of every even integer is divisible by 4".
ii. The only variable occurring in (10.173) is n. And it is bound. The sentence has no free variables.

iii. The sentence is a proposition.

iv.a The sentence is ***true***.

iv.b *Proof of (10.173):*

We want to prove the universal sentence (10.173), and we are going to do it using Rule $\forall_{prove}$.

Let n be an arbitrary integer.

We want to prove that $2|n \implies 4|n^2$. *And for that purpose we are going to use Rule* $\Longrightarrow_{prove}$.

Assume tbat $2|n$.
Then $(\exists k \in \mathbb{Z})n = 2k$.
Write $n = 2k$, $k \in \mathbb{Z}$. [Rule $\exists_{use}$]
Then $n^2 = (2k)^2 = 4k^2$.
Furthermore, $k^2 \in \mathbb{Z}$. [Reason: $k \in \mathbb{Z}$ and $(\forall k \in \mathbb{Z})k^2 \in \mathbb{Z}$.]
So k^2 is a witness for $(\exists k \in \mathbb{Z})n^2 = 4k$.
Therefore $(\exists k \in \mathbb{Z})n^2 = 4k$. [Rule $\exists_{prove}$]
So $4|n^2$. [Definition of "|"]

We have proved "$4|n^2$" assuming "$2|n$".

Hence $2|n \implies 4|n^2$. [Rule $\Longrightarrow_{prove}$]

We have proved "$2|n \implies 4|n^2$*" for arbitrary* $n \in \mathbb{Z}$.

Hence $\boxed{(\forall n \in \mathbb{Z})(2|n \implies 4|n^2)}$ [Rule $\forall_{prove}$]

Q.E.D.

Problem 50. For the sentence

$$(\forall n \in \mathbb{Z})2|n \implies 4|n^2\,, \tag{10.174}$$

perform exactly the same tasks i., ii., iii., and iv. (a. and b.) as in Problem 46.

Solution.

i. Sentence (10.174) says: "if every integer is even then n^2 is divisible by 4".

ii. The only variable occurring in (10.170ax) is n. This variable occurs in (10.174) three times; the first two occurrences are bound, and the third one is free.

iii. The sentence is a not a proposition.

iv.a,b Since the sentence is not a proposition, questions [iv.a] and [iv.b] do not apply.

Problem 51. For the sentence

$$(\forall n \in \mathbb{Z})(2|n \wedge 4|n^2), \tag{10.175}$$

perform exactly the same tasks i., ii., iii., and iv. (a. and b.) as in Problem 46.

Solution.

i. Sentence (10.175) says: "for every integer n, n is even and n^2 is divisible by 4".

ii. The only variable occurring in (10.173) is n. And it is bound. The sentence has no free variables.

iii. The sentence is a proposition.

iv.a The sentence is ***false***.

iv.b *Short proof that (10.175) is false:*

If "$(\forall n \in \mathbb{Z})(2|n \wedge 4|n^2)$" was true, then "$2|n \wedge 4|n^2$" would be true for every $n \in \mathbb{Z}$.

But "$2|n \wedge 4|n^2$" is false for $n = 1$.

So "$(\forall n \in \mathbb{Z})(2|n \wedge 4|n^2)$" is false.

Q.E.D.

Problem 52. For the sentence

$$(\forall n \in \mathbb{Z})(2|n \vee 4|n^2), \tag{10.176}$$

perform exactly the same tasks i., ii., iii., and iv. (a. and b.) as in Problem 46.

Solution.

i. Sentence (10.176) says: "for every integer n, n is even or n^2 is divisible by 4".

ii. The only variable occurring in (10.173) is n. And it is bound. The sentence has no free variables.

iii. The sentence is a proposition.

iv.a The sentence is ***false***.

iv.b *Short proof that (10.176) is false:*

If "$(\forall n \in \mathbb{Z})(2|n \vee 4|n^2)$" was true, then "$2|n \vee 4|n^2$" would be true for every $n \in \mathbb{Z}$.

But "$2|n \vee 4|n^2$" is false for $n = 1$.

So "$(\forall n \in \mathbb{Z})(2|n \vee 4|n^2)$" is false.

Q.E.D.

Problem 53. For the sentence

$$(\forall n \in \mathbb{Z})(2|n \iff 4|n^2), \tag{10.177}$$

perform exactly the same tasks i., ii., iii., and iv. (a. and b.) as in Problem 46.

Solution.

i. Sentence (10.173) says: "for every even integer n, n is even if and only if n^2 is divisible by 4".

ii. The only variable occurring in (10.173) is n. And it is bound. The sentence has no free variables.

iii. The sentence is a proposition.

iv.a The sentence is ***true***.

Short proof of (10.177):

We want to prove the universal sentence (10.177). According to Rule $\forall_{prove}$*, we can do this by proving "*$(2|n \iff 4|n^2)$*" for an arbitarry integer* n.

Let n be an arbitrary integer.

We want to prove the biconditional sentence "$2|n \iff 4|n^2$*". According to Rule* $\iff_{prove}$*, we can do this by proving the implications "*$2|n \implies 4|n^2$*" and "*$4|n^2 \implies 2|n$*".*

Short proof of "$2|n \implies 4|n^2$*".*

We have already proved that $(\forall n \in \mathbb{Z})(2|n \implies 4|n^2)$ in our solution of problem 49.

So "$2|n \implies 4|n^2$" follows by Rule $\forall_{use}$.

Short proof of "$4|n^2 \implies 2|n$*".*

Assume $4|n^2$.

We will prove "$2|n$*" by contradiction.*

Assume $\sim 2|n$.

Then n is odd.

So n^2 is odd. [Reason: the product of two odd integers is odd]

That is, $\sim 2|n^2$.

But $4|n^2$, so n^2 is even.

That is, $2|n^2$.

Hence $2|n^2 \wedge \sim 2|n^2$, which is a contradiction.

Since we have proved a contradiction assuming $\sim 2|n$, we can conclude that $2|n$.

Since we have proved $2|n$ assuming $4|n^2$, we can conclude, thanks to Rule $\implies_{prove}$, that $4|n^2 \implies 2|n$.

Since we have proved "$2|n \implies 4|n^2$" and "$4|n^2 \implies 2|n$", we can conclude, thanks to Rule $\iff_{prove}$, that $2|n \iff 4|n^2$.

We have proved "$2|n \iff 4|n^2$*" for arbitrary* $n \in \mathbb{Z}$.

Hence $\boxed{(\forall n \in \mathbb{Z})(2|n \iff 4|n^2)}$ [Rule $\forall_{prove}$]

Q.E.D.

11 A more detailed introduction to logic

11.1 First-order predicate calculus

The language most mathematicians use to talk about mathematical objects (numbers of various kinds, sets, functions, lists, points, lines, planes, curves of various kinds, spaces where we can do geometry, graphs, and millions of other things) is a ***first-order predicate calculus***.

So let us explain what this means.

- The language is a "predicate calculus" because we can use it to express predicates.

So let us review what "predicates" are.

11.1.1 Predicates

Remember that

> A ***predicate*** is a sentence[a] involving one or more (or zero) variables, in such a way that the sentence has a definite truth value[b] for each choice of values of the variables.
>
> [a] "Sentence" means the same as "statement", or "assertion".
>
> [b] The ***truth value*** of a sentence is "true" if the sentence is true and "false" if the sentence is false.

For example:

- "Alice likes Mark" is a zero-variables predicate. It is either true or false.
- "x likes Mark" is a one-variable predicate. It is true or false depending on who x is. For example, suppose that Alice likes Mark but Andrew does not like Mark. Then "x likes Mark" is true when x =Alice but "x likes Mark" is false when x =Andrew.

  If we call this predicate $P(x)$, then $P(Alice)$ is true and $P(Andrew)$ is false.
- "x likes y" is a two-variables predicate. It is true or false depending on who x and y are. For example, suppose that Alice likes Mark, Andrew does not like Mark, Andrew likes Alice, and Mark does not like Andrew. Then "x likes y" is true when x =Alice and y =Mark,

and when x =Andrew and y =Alice, but "x likes y" is false when x =Andrew and y =Mark.

If we call this predicate $P(x, y)$, then $P(Alice, Mark)$ is true but on the other hand $P(Mark, Andrew)$ is false.

- If S is the set of all people, then "$(\forall x \in S)x$ likes y" says "everybody likes y". This is a one-variable predicate. We could call this predicate $Q(y)$, and then we could define $Q(y)$ as follows:

$$\text{if } y \in S \text{ then } Q(y) \text{ means } (\forall x \in S)P(x, y)\,, \tag{11.178}$$

or, in purely formal language:

$$(\forall y \in S)\Big(Q(y) \iff (\forall x \in S)P(x, y)\Big) \tag{11.179}$$

- "x likes y more than x likes z" is a three-variables predicate.
- "$2 + 2 = 4$" and "$2 + 2 = 5$" are zero-variables predicates. They are either true or false. (And, of course, "$2+2 = 4$" is true and "$2+2 = 5$" is false.)
- "$x > 0$" and "$2|n$" are one-variable predicates. They are true or false depending on who x (or n) is. For example, "$x > 0$" is true $x = 3$ but is false for $x = -5$. And "$2|n$" is true for $n = 4$ but is false for $n = 5$.
- "$x > y$" and "$m|n$" are two-variables (i.e., binary) predicates. They are true or false depending on who x and y (or m and n) are. For example, "the sentence $x > y$" is true for $x = 5$ and $y = 4$, but is false for $x = 5$ and $y = 6$. And "$m|n$" is true for $m = 3$ and $y = 6$, but is not true for $m = 3$ and $y = 7$.
- "$x+y = z$", "$x+y > z$", and "$n|m+q^2$" are three-variables predicates. The predicate "$x + y = z$" is,true for $x = 2$, $y = 3$ and $z = 5$, but is false for $x = 2$, $y = 3$ and $z = 4$. The predicate "$x + y > z$" is true for $x = 2$, $y = 3$ and $z = 4$. but is false for $x = 2$, $y = 3$ and $z = 5$. The predicate "$n|m + q^2$" is true for $n = 5$, $m = 9$, and $q = 6$, but is false $n = 5$, $m = 7$, and $q = 6$.
- "$x + 2y^2 - z > u$" and "$a = bq + r$ and $0 \leq r < |b|$" are four-variables predicates. The predicate "$x + 2y^2 - z > u$" is true for $x = 2$, $y = 4$. $z = 3$, $u = 4$, but is false for $x = 2$, $y = 1$. $z = 3$, $u = 3$, The predicate "$a = bq+r$ and $0 \leq r < |b|$" is true for $a = 23$, $b = 5$, $q = 4$ and $r = 3$, but is false for $a = 23$, $b = 5$, $q = 4$ and $r = 2$.

11.2 Free and bound variables, quantifiers, and the number of variables of a predicate

As was explained in the previous section, in a predicate such as "$x > y$", the variables x, y are ***free variables***, that is, variables that are free to be given any value we want. We can plug in values for x and y, and for each choice of values the resulting sentence has a definite truth value, that is, is true or false.

> You should think of a predicate as a ***processing device***, with several "input channels". The input channels are the ***open variables***. Each input channel is ***open***, in the sense that the entrance to the channel is open so you can can put things in, or ***free***, in the sense that we are free to put things in there. Once you have put in a value for, say, the variable x, then x is no longer open: it becomes ***closed***, or ***bound***.
> Once you have put in values in all the input channels, the device processes these inputs, and produces a answer: true, or false.

If, on the other hand, the predicate "$x > y$" appears in a text after a statement such as

$$\text{Let } x = 5,\ y = 3.$$

then the variables x and y are no longer free: they are ***bound variables***[60], because they are "attached" to particular values.

We now look at another, very important way to turn free variables into bound variables.

Let us consider, for example, the predicates

$$(\forall y \in \mathbb{R})x + 2y^2 - z > u \tag{11.180}$$

and

$$(\exists q \in \mathbb{Z})(\exists r \in \mathbb{Z})(a = bq + r\,, \text{ and } 0 \leq r < |b|)\,. \tag{11.181}$$

You may think that these are four-variables predicates, because each one of them contains four variables. (Predicate (11.180) contains the variables x, y, z and u. Predicate (11.181) contains the variables a, b, a and r.)

But this is not right:

[60] Bound variables are also called ***closed variables***, because they are not open: the "input channel" through which we can input values for the variables is closed.

(11.180) is a three-variables predicate, and ***(11.181) is two-variables predicate.***.

Let me explain.

11.2.1 An example: a predicate with three free variables and one bound variable

We first look at the predicate

$$(\forall y \in \mathbb{R})x + 2y^2 - z > u\,. \tag{11.182}$$

- The predicate (11.182) is built from the predicate "$x+2y^2-z > u$" by ***quantifying*** it, i.e., putting a universal quantifier $(\forall y \in \mathbb{R})$ in front.
- The unquantified predicate "$x + 2y^2 - z > u$" contains the variables x, y, z, u. These are four open variables.
- So, if you are asked the "truth question"

  $$\text{Is} \quad \text{"}x + 2y^2 - z > u\text{"} \quad \text{true or false?}$$

  then you have to reply with a question of your own:

  $$\text{Who are } x,\ y,\ z \text{ and } u?$$

- But in the quantified predicate (11.182) ***the variable y is quantified***.
- So, if you are asked the "truth question"

  $$\text{Is} \quad \text{"}(\forall y \in \mathbb{R})x + 2y^2 - z > u\text{"} \quad \text{true or false?}$$

  then you have to reply with the question:

  $$\text{Who are } x,\ z \text{ and } u?$$

- In the predicate "$x + 2y^2 - z > u$", the four variables x, y, z and u are open variables, that is, "slots", or "input channels", where you can put in (or "plug in") values for each of the variables.

- When you fill in the four slots by plugging in values for the variables, you get a ***proposition***, i.e., a sentence that has a definite truth value.

A proposition is a sentence with no open variables

So a proposition is just true or false, whereas a predicate with open variables is true or false depending on the values of the variables.
Example:

1. The sentence "$m \geq n$" has two open variables. It is true if, for example, $m = 3$ and $n = 1$, and it is false if, for example, $m = 3$ and $n = 4$.

2. The sentence "$(\forall m \in \mathbb{N})m \geq n$" is true if, for example, $n = 1$, and it is false if, for example, $n = 2$. So this sentence has one open variable, namely, n.

3. The sentences

$$(\exists n \in \mathbb{N})(\forall m \in \mathbb{N})m \geq n$$

and

$$(\forall n \in \mathbb{N})(\forall m \in \mathbb{N})m \geq n$$

do not have any open variables. So they are propositions. The first one is true. (Reason: Take $n = 1$. Then for arbitrary $m \in \mathbb{N}$ $m \geq 1$.) The second one is false. (Reason: Take $m = 1$, $n = 2$. Then it is not true that $m \geq n$.)

- So, for example, if you plug in the values $x = 2$, $y = 4$, $z = 3$, $u = 4$, into the sentence

$$x + 2y^2 - z > u$$

you get the proposition

$$19 > 4, ,$$

which is true.

- But in the quantified predicate "$(\forall y \in \mathbb{R})x + 2y^2 - z > u$", there is no y-slot. The three variables x, z and u are open variables, that is, slots or input channels where you can put in values. But y is not an open variable.

- When you fill in the slots by plugging in values for the three open variables, you get a proposition.

- So, for example, if you plug in the values $x = 2$, $z = -3$, $u = 4$, into the sentence

$$(\forall y \in \mathbb{R})x + 2y^2 - z > u$$

then you get the sentence

$$(\forall y \in \mathbb{R})2 + 2y^2 + 3 > 4$$

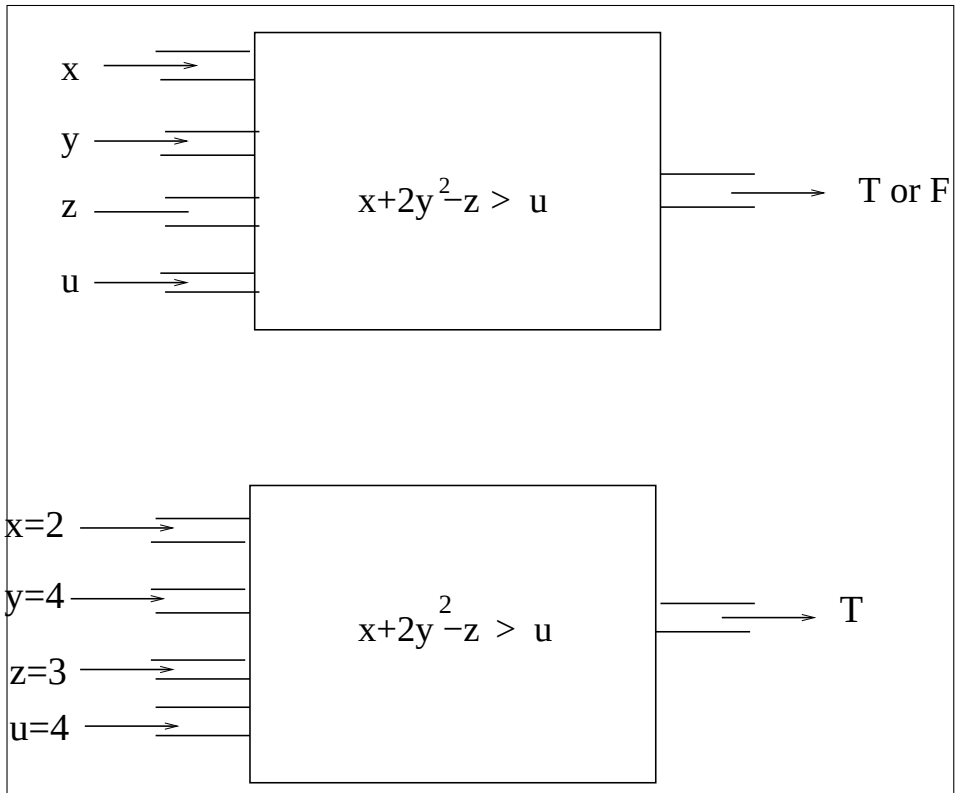
which is equivalent to the sentence

$$(\forall y \in \mathbb{R})2y^2 + 5 > 4\,.$$

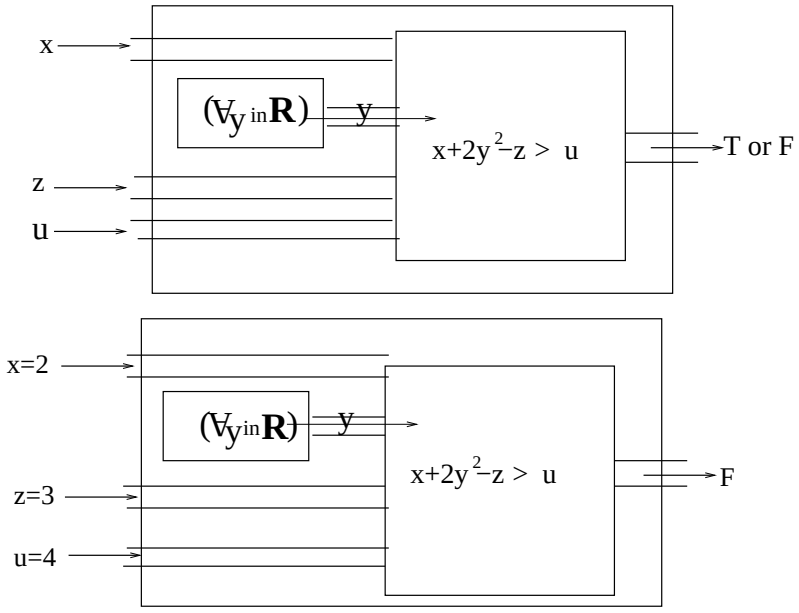
And this sentence is true. (Proof: Let $y \in \mathbb{R}$ be arbitrary. Then $2y^2 \geq 0$. But $5 > 4$. So $2y^2 + 5 > 4$. Hence "$2y^2 + 5 > 0$" is true for arbitrary $y \in \mathbb{R}$. Therefore "$(\forall y \in \mathbb{R})2y^2 + 5 > 4$" is true.)

- The key point here is that ***the sentence "$(\forall y \in \mathbb{R})x + 2y^2 - z > u$" does not have a y-slot where you can plug in a value of y.*** That's because ***the sentence itself decides which value or values of y to plug in.*** The quantifier $(\forall y \in \mathbb{R})$ says: "let y be an arbitrary real number". And then, with the values of x, z and u supplied by you, the truth value of the resulting sentence is determined. ***There is no need to ask "who is y?"***

Another way to see this is as follows: when you universally quantify a sentence by putting in front of it the universal quantifier "$(\forall y \in \mathbb{R})$", this adds to the sentence a "generator of y-values", that is, a new component that tells the sentence what value of y to use. More precisely, the universal quantifier "$(\forall y \in \mathbb{R})$" says "Let y be an arbitrary real number". And this ***closes*** the y-input channel, so that it is no longer possible to plug a y-value into the sentence from outside.



The sentence $\boxed{x + 2y^2 - z > u}$ is a processing device that has four input channels: the x-channel, the y-channel, the z-channel, and the u-channel. When values for the four variables are inputted into the sentence, the sentence produces a truth value. The four variables x, y, z, u are ***open***, or ***free***. They are open, because the input channels are open so that values of the variables can be put into them. They are free, because the variables are not tied to any particular value.



The quantified sentence $\boxed{(\forall y \in \mathbb{R})x + 2y^2 - z > u}$ is a combination of two interconnected processing units: the original unquantified sentence "$x+2y^2-z>u$", and the quantifier "$(\forall y \in \mathbb{R})$". The quantifier generates a value for the quantified variable y (by saying "let y be an arbitrary real number") and, by doing so, it ***closes*** the y input channel, so that y is no longer free; we cannot choose a value for y and plug it in. The other three channels remain open. So in this sentence x, z and u are open variables. but y is ***closed***, or ***bound***.

The other three letter variables (x, z and u) remain open. So we can plug in values for them in order to obtain propositions that have a definite truth value.

Summarizing:

- Even though the predicate "$(\forall y \in \mathbb{R})x + 2y^2 - z > u$" appears to contain four letter variables, only three of these variables (x, z and u) are open. The other variable, y, is ***bound***, or ***closed.***
- This means that the predicate "$(\forall y \in \mathbb{R})x + 2y^2 - z > u$" is a ***three variables***, or ***three arguments***, predicate. Therefore:
  - For each choice of values for x, z and u, the predicate becomes a proposition, i.e. a sentence with a definite truth value.

  - If we want to give a name to this predicate, then we can of course call it P, but if we want to indicate the names of the free variables, we should call it $P(x, z, u)$.
  - But ***we must not call it*** $P(x, y, z, u)$, because if we give it such a name we would erroneously be suggesting that this predicate has a "y-channel" where we can input values for the variable y.

- For example, "$(\forall y \in \mathbb{R})x + 2y^2 - z > u$" is true for $x = 4$, $z = 2$, $u = 1$. (Proof: We want to prove that $(\forall y \in \mathbb{R})4 + 2y^2 - 2 > 1$, that is, that $(\forall y \in \mathbb{R})2 + 2y^2 > 1$. Let $y \in \mathbb{R}$ be arbitrary. Then $y^2 \geq 0$, so $2y^2 \geq 0$, so $2 + 2y^2 \geq 2$, and $2 > 1$, so $2 + 2y^2 > 1$. Since "$2+2y^2 > 1$" has been proved to be true for arbitrary real y, it follows that $(\forall y \in \mathbb{R})2 + 2y^2 > 1$. Q.E.D.)

- The predicate "$(\forall y \in \mathbb{R})x + 2y^2 - z > u$" is false for $x = 4$, $z = 2$, $u = 8$. (Proof: We want to prove that "$(\forall y \in \mathbb{R})4 + 2y^2 - 2 > 1$" is not true, i.e., that "$(\forall y \in \mathbb{R})2 + 2y^2 > 8$" is not true. Take $y = 0$. Then "$2 + 2y^2 > 8$" is not true, because "$2 + 0 > 8$" is not true. So "$(\forall y \in \mathbb{R})4 + 2y^2 - 2 > 1$" is not true.Q.E.D.)

- The "truth question", i.e., the extra question we need to ask is order to be able to tell if "$(\forall y \in \mathbb{R})x + 2y^2 - z > u$" is true or false, is the question: ***"who are*** x***,*** z ***and*** u***?"***

- ***in order to have enough information to determine if the sentence "***$(\forall y \in \mathbb{R})x+2y^2-z > u$***" is true or false, we do not have to ask "who is*** y***?", because once you are given the values of*** x***,*** z ***and*** u***, the quantified sentence itself determines if it is true or false, because it is up to the sentence to decide if it's true for all*** y ***or not, and it's not up to you to choose a value for*** y***.***

11.2.2 A second example: a predicate with two free variables and two bound variables

We now look at the predicate

$$(\exists q \in \mathbb{Z})(\exists r \in \mathbb{Z})(a = bq + r \wedge 0 \leq r < |b|)\,. \tag{11.183}$$

As I said before, on page 183, ***(11.181) is a two-variables predicate.***.

- Predicate (11.183) contains the variables a, b, q and r. But q ***and*** r ***are quantified***. So, if you are asked the "truth question"

Is "$(\exists q \in \mathbb{Z})(\exists r \in \mathbb{Z})(a = bq + r \wedge 0 \leq r < |b|)$" true or false?

then you have to reply with a question of your own:

Who are a and b?

The variables a and b in (11.183) are "slots", or "input channels", where you can put in (or "plug in") a value for each of the variables, and then you get a proposition.

- So, for example, if you plug in the values $a = 23$, $b = 11$, into the sentence
$$(\exists q \in \mathbb{Z})(\exists r \in \mathbb{Z})(a = bq + r \wedge 0 \leq r < |b|)$$
then you get the sentence
$$(\exists q \in \mathbb{Z})(\exists r \in \mathbb{Z})(23 = 11q + r \wedge 0 \leq r < 11)\,.$$
And this sentence is true. (Proof: To prove an existential statement we use rule $\exists_{use}$: we exhibit values of q and r for which the proposition "$23 = 11q + r \wedge 0 \leq r < 11$" is true. Take $q = 2$, $r = 1$. Then $23 = 11q + r$ and $0 \leq r < 11$. Hence "$23 = 11q + r \wedge 0 \leq r < 11$" is true for some $q, r \in \mathbb{Z}$. Therefore "$(\exists q \in \mathbb{Z})(\exists r \in \mathbb{Z})(23 = 11q + r \wedge 0 \leq r < 11$" is true.)

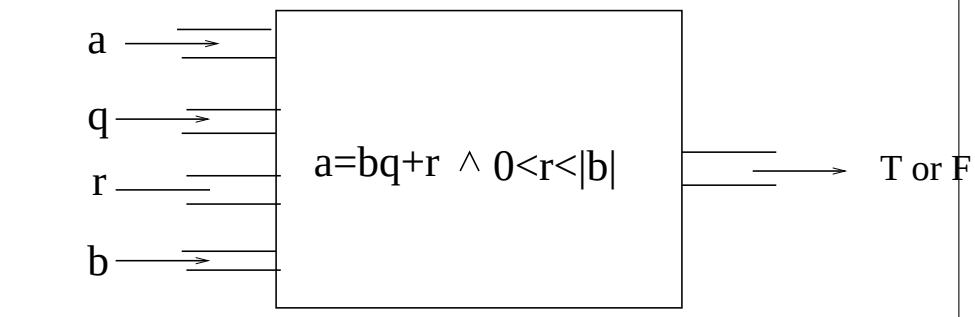
- The key point here is that ***the sentence***
$$(\exists q \in \mathbb{Z})(\exists r \in \mathbb{Z})(23 = 11q + r \wedge 0 \leq r < 11)$$
does not have a q-slot or an r-slot where you can plug in values for q and r. That's because ***the sentence itself decides which value or values of q and r to plug in.*** The sentence itself[61] decides which values of q and r it has to look at, and then, with the values of a and b supplied by you, the truth value of the resulting sentence is determined.

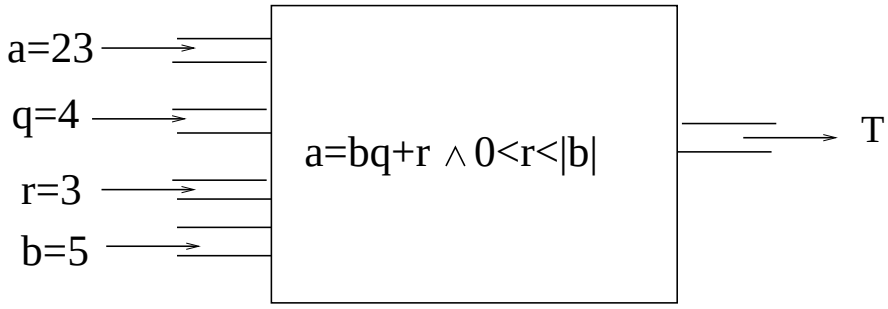
[61] Remember: you must think of a sentence as a processing device. The unquantified sentence "$a = bq + r \wedge 0 \leq r < |b|$" does the following: once it has been fed values for a, b, q and r, it finds out if "$a = bq + r \wedge 0 \leq r < |b|$" is true or not; if it is true is says "yes"; if it is false it says "no". The quantified sentence "$(\exists q \in \mathbb{Z})(\exists r \in \mathbb{Z})(23 = 11q + r \wedge 0 \leq r < 11)$" does a much more complicated job: once it has been fed values for a and b, the sentence looks at all possible values of q and r, and sees whether it can find one choice of values of q and r for which "$23 = 11q + r \wedge 0 \leq r < 11$" is true; and then, if it find such values, it says "yes"; and if it cannot find any values of q and r for which "$23 = 11q + r \wedge 0 \leq r < 11$" is true, it says "no".

- Another way to see this is as follows: the sentence "$a = bq + r \wedge 0 \leq r < |b|$" has four input channels that are open, or free, so you can put into each channel a value of the corresponding variable.

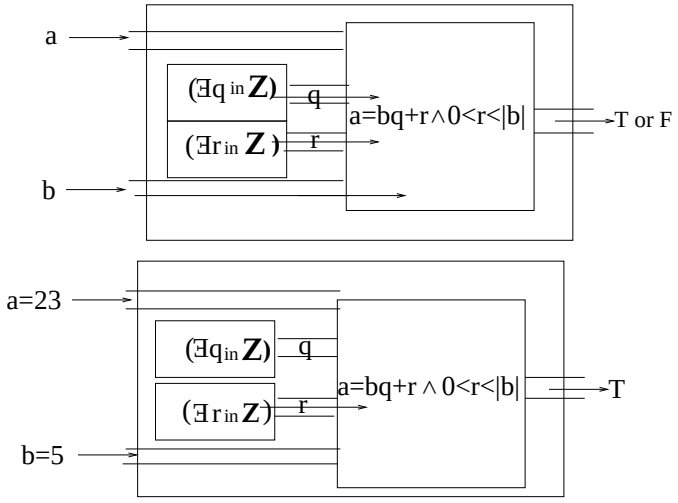
  But when you existentially quantify the sentence twice by putting in front of it the two existential quantifiera "$(\exists q \in \mathbb{Z})$" and "$(\exists r \in \mathbb{Z})$", this adds to the sentence a "generator of q-values" and a "generator of r-values", that is, two new components that tell the sentence what values of q and r to look at. More precisely, the existential quantifiers "$(\exists q \in \mathbb{R})$" and "$(\exists r \in \mathbb{R})$" do the following:

  - They look for a q-value and an r-value that make the sentence "$a = bq + r \wedge 0 \leq r < |b|$" true.
  - If they find such values, then they send to the sentence the message "yes, we have found values that make you true", and then the sentence produces the final verdict "yes, true".
  - If they do not find such values, then they send to the sentence the message "no, we have not found values that make you true", and then the sentence produces the final verdict "no, not true".





The sentence $\boxed{a = bq + r \wedge 0 \leq r < |b|}$ is a processing device that has four input channels:the a-channel, the q-channel, the r-channel, and the b-channel. When values for the four variables are inputted into the sentence, the sentence produces a truth value. The four variables a, q, r, b are ***open***, or ***free***. They are open, because the input channels are open so that values of the variables can be put into them. They are free, because the variables are not tied to any particular value.



The quantified sentence $\boxed{(\exists q \in \mathbb{Z})(\exists r \in \mathbb{Z})(a = bq + r \wedge 0 \leq r < |b|)}$ is a combination of three interconnected processing units: the original unquantified sentence "$a = bq + r \wedge 0 \leq r < |b|$)", and the two quantifiers "$(\exists q \in \mathbb{Z}$", "$(\exists r \in \mathbb{Z}$". The quantifiers generate values for the quantified variablea q, r. (They look for values of q, r that will make "$a = bq + r \wedge 0 \leq r < |b|$)" true. If they find them, then they send one pair of such values to the main processing unit "$a = bq + r \wedge 0 \leq r < |b|$)", which then says "yes, true". If they do not find them, then they send some values to the main processing unit, but these values will not work, so the main processing unit wil say "no, not true".) By doing so, the quantifiers ***close*** the q and r input channels, so that q and f are no longer free; we cannot choose values for q and r and plug them in. The other two channels remain open. So in this sentence a and b are open variables. but q and r are ***closed***, or ***bound***.

The other two letter variables (a and b) remain open. So we can plug in values for them in order to obtain propositions that have a definite truth value.

Summarizing:

- Even though the predicate

$$(\exists q \in \mathbb{Z})(\exists r \in \mathbb{Z})(a = bq + r \wedge 0 \leq r < |b|)$$

  appears to contain four letter variables, only two of these variables (a and b) are open. The other variables, q and r, are ***bound***, or ***closed.***

- This means that the predicate

$$(\exists q \in \mathbb{Z})(\exists r \in \mathbb{Z})(a = bq + r \wedge 0 \leq r < |b|)$$

  is a ***two variables***, or ***two arguments***, predicate. Therefore:

  - For each choice of values for a and b, the predicate becomes a proposition, i.e. a sentence with a definite truth value. (And the Division Theorem tells us that the truth value is "true" for all choices of integers a and b such that $b \neq 0$, that is, that the proposition[62]

$$(\forall a \in \mathbb{Z})(\forall b \in \mathbb{Z})\Big(b \neq 0 \Longrightarrow (\exists q \in \mathbb{Z})(\exists r \in \mathbb{Z})(a = bq + r \wedge 0 \leq r < |b|)\Big) \quad (11.184)$$

    is true.
  - Suppose we want to give a name to the two-variables predicate

$$(\exists q \in \mathbb{Z})(\exists r \in \mathbb{Z})(a = bq + r \wedge 0 \leq r < |b|)\,.$$

    We can, of course, call it P. But if we want to indicate the names of the free variables, we should call it $P(a, b)$.
  - But ***we must not call it*** $P(a, b, q, r)$, because if we give it such a name we would erroneously be suggesting that this predicate has a "q-channel" and an "r-channel", where we can input values for the variables q, r.

- The "truth question", i.e., the extra question we need to ask is order to be able to tell if "$(\exists q \in \mathbb{Z})(\exists r \in \mathbb{Z})(a = bq + r \wedge 0 \leq r < |b|)$" is true or false, is the question: ***"who are*** a ***and*** b***?"***

- ***in order to have enough information to determine if the sentence "***$(\exists q \in \mathbb{Z})(\exists r \in \mathbb{Z})(a = bq + r \wedge 0 \leq r < |b|)$***" is true or false, we do not have to ask "who are*** q ***and*** r***?", because once you are given the values of*** a ***and*** b***, the quantified sentence itself determines if it is true or false, because it is up to the sentence to decide if the required values of*** q ***and*** r ***exists or not, and it's not up to you to choose valuea for*** q ***and*** r***.***

[62] Notice that (11.184) is a proposition, i.e., a predicate with no open variables at all (or, if you prefer, with zero open variables), because in (11.184) all four variables that occur are quantified, so a, b, q and r are closed variables. For the sentence (11.184), if you are asked "is this true", you do not need to ask any "truth question", because you do not need values of any variables to determine if the sentence is true.

11.2.3 Another example, illustrating the fact that only open variables really matter

Some natural numbers are products of two prime numbers; for example, $4 = 2 \times 2$, $6 = 2 \times 3$, $35 = 5 \times 7$, and so on, Other natural numbers are not products of two prime numbers; for example, $18 = 2 \times 3 \times 3$, and the Fundamental Theorem of Arithmetic tells us that that there is no other way to write 18 as a product of primes, so in particular 18 is not the product of two primes.

So we can consider the predicate "n is a product of two prime numbers". And we can call this predicate $A(n)$. (We could just have called is "A", but we choose the name "$A(n)$" to emphasize the fact that this predicate has the open variable n.) Then, according to the conventions we made before about naming predicates, $A(6)$ is the proposition "6 is the product of two primes", and $A(7)$ is the proposition "7 is the product of two primes", so $A(6)$ is true, and $A(7)$ is false.

You can think of the predicate $A(n)$ as a "black box": you input a value of n, the predicate does some work, and produces an answer: "true" or "false". (For example, for $n = 6$ $A(n)$ is true, and for $n = 7$ $A(n)$ is false.)

But we can also look inside the box, and analyze in more detail how this predicate works. That is, we can observe that $A(n)$ says that

$$\text{There exist prime numbers } p, q \text{ such that } n = pq.$$

So now our predicate has three variables, p, q, and n!

How come? Has the number of variables of $A(n)$ suddenly changed? Has $A(n)$ become a three-variables predicate? You may think so, because now $A(n)$ seems to have three variables: p, q and n.

But the answer is: ***No!*** $A(n)$ ***is still a one-variable predicate!*** The variables p and q are ***bound***, because they are quantified. Precisely, $A(n)$ says, in semiformal (almost formal) language:

$$(\exists p \in \mathbb{N})(\exists q \in \mathbb{N})(p \text{ is prime} \wedge q \text{ is prime} \wedge n = pq)\,. \tag{11.185}$$

So, even though $A(n)$ appears to have three variables, namely, p, q and n, two of them are ***internal variables***[63], within the sentence (11.185). The sentence itself generates the values of p and q that it needs in order to answer

[63] If you think of the sentence "$(\exists p \in \mathbb{N})(\exists q \in \mathbb{N})(p \text{ is prime})$" as a processing unit, you will see that it has two existential quantifiers that generate values of p and q. But outside the processing unit all one sees is that certain values of n are fed in and certain 'true"s and "false"s come out. The variables p and q are part of the internal operation of the device.

its true-false question, and when the sentence ends p and q are free variables again. And, in particular, ***outside the sentence***

$$(\exists p \in \mathbb{N})(\exists q \in \mathbb{N})(p \text{ is prime} \wedge q \text{ is prime} \wedge n = pq)$$

the variables p and q have no value.

Another way to see that p and q have no value, is to observe that $A(n)$ can equally well be written as

$$(\exists x \in \mathbb{N})(\exists y \in \mathbb{N})(x \text{ is prime} \wedge y \text{ is prime} \wedge n = xy)\,, \tag{11.186}$$

or as

$$(\exists u \in \mathbb{N})(\exists v \in \mathbb{N})(u \text{ is prime} \wedge v \text{ is prime} \wedge n = uv)\,. \tag{11.187}$$

Sentences (11.185), (11.186), and (11.187) say exactly the same thing. The only difference is in the names of the variables that, inside the box, the sentence uses to process the inputs and produce an output.

From outside the box, we do not see these variables. ***That's why the letters p, q in (11.185), as well as the letters x, y in (11.186), and the letters u, v in (11.187), are internal variables, that have no value outside the sentence.***

And this is not the end of the story. "p is prime" is itself a complex predicate. In fact, "p is prine" stands for

$$p > 1 \wedge (\forall k \in \mathbb{N})\Big(k|p \Longrightarrow (k = 1 \vee k = p)\Big)\,. \tag{11.188}$$

This means that $A(n)$ can also be written as

$$(\exists p \in \mathbb{N})(\exists q \in \mathbb{N})\Bigg(\Bigg(p > 1 \wedge (\forall k \in \mathbb{N})\Big(k|p \Longrightarrow (k = 1 \vee k = p)\Big)\Bigg)$$
$$\wedge\Bigg(q > 1 \wedge (\forall k \in \mathbb{N})\Big(k|q \Longrightarrow (k = 1 \vee k = q)\Big)\Bigg) \wedge n = pq\Bigg)\,. \tag{11.189}$$

Now one may think that $A(n)$ is a four-variables predicate, because it involves the variables n, p, q and k. But by now you know better: the new variable k is also bound, so the only open variable in (11.189)) is still n. That means that ***even if you write it in the form (11.189), $A(n)$ is still a one-variable predicate.***

Actually, the story doesn't end here either. "$k|p$" is also a complex preddcate with an internal structure of its own: is stands for "$(\exists j \in \mathbb{Z})p =$

kj". So, if we substitute this for "$k|p$" in (11.189), we get an even more detalied version of $A(n)$, namely,

$$\begin{aligned}&(\exists p \in \mathbb{N})(\exists q \in \mathbb{N})\\ &\quad\Bigg(\bigg(p > 1 \wedge (\forall k \in \mathbb{N})\Big((\exists j \in \mathbb{Z})p = kj \Longrightarrow (k = 1 \vee k = p)\Big)\bigg)\\ &\quad\wedge\bigg(q > 1 \wedge (\forall k \in \mathbb{N})\Big((\exists j \in \mathbb{Z})q = kj \Longrightarrow (k = 1 \vee k = q)\Big)\bigg)\\ &\quad\wedge n = pq\Bigg)\,. \end{aligned} \tag{11.190}$$

Now $A(n)$ apears to involve five variables: n, p, q, k and j. But this time you will have no problem figuring out that $A(n)$ ***is still a one-variable predicate, because the only open variable in (11.190) is still*** n***, and all the other variables are bound***.

Problem 54. Draw a diagram of the sentence (11.190) as a processing unit, similar to the diagrams that appear on pages 187 and 192.

Make sure that your diagram shows that there is only only one input channel. □

11.2.4 Dummy variables

So far, we have seen that variables that appear in a sentence but are quantified are "internal variables", or "closed variables", or "bound variables". If you think of a sentence as a "processing unit", or "processing device", that takes in certain inputs and produces "true-false" outputs, then the closed (or bound, or internal) variables are variables that the sentence itself generates and uses to do its processing work. So the sentence does not need to be fed the values of these variables, and does not produce values of those variables that an outside obsevrer can see.

There is another way in which a variable appearing in a sentence can be a closed (or bound, or internal) variable. The sentence may contain a part that generates values of some variable in order to do a computation.

Consider. for example, the sentence

$$\sum_{k=1}^{n}(a + r^k) = b\,, \tag{11.191}$$

This sentence contains five letter variables, namely, a, r, b, k, and n.

Which ones of these five variables are open?

The best way to answer this question is by thinking of (11.191) as a processing device, opening it up to look into its internal structure, and figuring out what the device does.

Suppose you ask the device the truth question:

$$\text{Is it true that } \sum_{k=1}^{n}(a+r^k)=b?$$

Then the device will not know what to do, because in order to get started the device needs to be given the values of a, b, r, and n. (Maybe we should think of (11.191) as an inteligent device, that can ask questions. Then if you ask the truth question, the device will answer with a question: ***who are*** a***,*** b***,*** r ***and*** n***?***)

Suppose you do feed the device by inputting values for a, b,r and n. Then the device will do the following:

1. First, the CPU (central procssing unit) will report to the summation component Σ—that is, the component that computes the summation $\sum_{k=1}^{n}(a+r^k)$—the values of a, b, r and n that it has received from you.

2. Then Σ will perform the following calculation:

   (a) First, it will write the list of all values of k, from 1 to n. (This is something it can do, because it knows who n is, since it has received this information from the CPU.)

   (b) Then it will compute $a+r^k$ for each of the values of k in the list. (Again, Σ knows how to do this, because it knows who a and r are.)

   (c) Then it will take all those values of $a+r^k$ that it has computed, and add them.

   (d) Finally, it will report the result to the CPU. (Maybe, in order to facilitate communication between Σ and the CPU, they will introduce letter variables. For example, they may decide to call the result of the summation s, and then Σ will report the value of s to the CPU. But we need not concern ourselves with the variable s, because that's an internal variable used within the device for the various parts to communicate with each other.)

3. The CPU will then compare the result reported by the summation unit with b, and determine if they are equal.

4. If they are equal, the CPU will report to you the answer "true".

5. If they are not equal, the CPU will report to you the answer "false".

The main point of this is that k ***is an internal variable used by the sentence to perform its calculation. The values of*** k ***are generated by the sentence itself. So the sentence need not be given the value of*** k. And that's why

1. If asked the truth question, the sentence will ask "who are a, b, r and n"".

2. The sentence will not ask "who is k?", because ***the sentence itself generates the values of*** k ***it needs***.

3. k ***is not an open variable in (11.191)***

4. The open variables of (11.191) are a, b, r and n.

Let's just look at one more example. Let us analyze the following four sentences

$$(\forall n \in \mathbb{N})\Big((\exists m \in \mathbb{N}) \sum_{k=1}^{m} k^3 = n \Longrightarrow (\exists p \in \mathbb{N}) n = p^2\Big)\,, \tag{11.192}$$

$$(\forall n \in \mathbb{N})\Big((\exists m \in \mathbb{N}) \sum_{k=1}^{m} k^3 = n \Longrightarrow (\exists p \in \mathbb{N}) n = p^3\Big)\,, \tag{11.193}$$

$$(\forall n \in \mathbb{N})(\exists m \in \mathbb{N})\Big(\sum_{k=1}^{m} k^3 = n \Longrightarrow (\exists p \in \mathbb{N}) n = p^2\Big) \tag{11.194}$$

and

$$(\forall n \in \mathbb{N})(\exists m \in \mathbb{N})\Big(\sum_{k=1}^{m} k^3 = n \Longrightarrow (\exists p \in \mathbb{N}) n = p^3\Big) \tag{11.195}$$

Each of these sentences contains four variables, namely, n, m, k, and p.

And I am sure that this time you can see right away what is going on: ***all four variables are closed***. Three of them (n, m, and p) are quantified. and the variable k is also closed because the sentence itself generates the values of k that it needs to perfom its calculations.

So ***the sentences (11.192), (11.193), (11.194), and (11.195), are propositions***.

And then of course each of the sentences is true or false. Which leads me to a natural question, that I will ask you to answer.

Problem 55. Which of the propositions (11.192), (11.193), (11.194), (11.195), are true, and which ones are false?

NOTE: All these propositions are of the form $(\forall n \in \mathbb{N})P(n)$, where $P(n)$ is a one-variable predicate having n as the open variable.

If you want to prove that a sentence of this form is true, then you need a reasoned argument, starting with "Let n be an arbitrary natural number." (You may also try a proof by induction, but in this case I would not recommend that.) If you want to prove that it is false, then you need a counterexample, i.e., an example of an n for which the one-variable sentence $P(n)$ is false.

HINT: The answer to this problem is actually very easy. All you have to do is use the result of one of your earlier homework problems. (I can narrow this down a bit further: *it's one of the problems in the third set of lecture notes*.) Using this, plus a little bit of logic (for example, truth values of implications), each of the four propositions should just require a couple of lines on your part.) □

A variable such as the k in $\sum_{k=1}^{n} t(k)$ (where $t(k)$ is some expression containing k, such as k, or k^2, or r^k, or $a+r^k$), is called a "dummy variable".

Let us define this term precisely. (The definition I am about to give is taken from Wolfram MathWorld.)

Definition 16. *A dummy variable is* ***a variable that appears in a calculation only as a placeholder and which disappears completely in the final result.*** □

And ***every dummy variable is bounded***.

Example 50. Naturally, summations are not the only type of expressions where some of the variables are bound variables

Examples of dummy variables are:

1. the k in a summation such as $\sum_{k=1}^{n} t(k)$,
2. the k in a product such as $\prod_{k=1}^{n} t(k)$,
3. the k in the name of a list, such as $(p_k)_{k=1}^{n}$,
4. the x in the name $\{x : P(x)\}$ of a set,
5. the x in an integral such as $\int_a^b f(x)dx$.
6. the x in a limit such as $\lim_{x\to a} f(x)$. □

Example 51. Let us look at the sentence

$$(\exists a \in \mathbb{R})(\exists b \in \mathbb{R})\Big(\{u \in \mathbb{R} : a \leq u \leq b\} \neq \emptyset \wedge \int_a^b x^2 dx = c\Big). \quad (11.196)$$

This sentence contains the letter variables a, b, u, x, and c.

Of these five letters, four are bound variables:

1. the variables a and b are bound because they are quantified;
2. the variable u is bound because it is a dummy variable, used as part of the name $\{u \in \mathbb{R} : a \leq u \leq b\}$ of a set;
3. the variable x is bound because it is a dummy variable, used as a variable of integration.

It follows from this analysis that

1. ***Sentence (11.196) defines a one-variable predicate.***
2. ***The open variable in sentence (11.196) is*** c.
3. If you think of sentence (11.196) as a processing device, then this device will take values of c as inputs, and produce a true-false answer as output.
4. If you ask the "truth question" ***is (11.196) true?***, then the device (11.196) cannot answer because it does not know who c is. So the device will answer your question with another question: ***who is*** c***?***
5. But, in order to be able to answer the truth question, the device does not need to ask "who is a?", or "who is b?" or "who is u?", or "who is x?". The device itself will generate the values of a, b, u and x it needs, and these values will be part of the calculations that (11.196) performs, and will not be seen by the outside world.

11.2.5 How to tell if a variable is dummy

Here are two ways to see that a variable is dummy.

1. The variable is dummy if "it isn't really there", in the sense that we can eliminate it completely. For example,

(a) The set $\{u \in \mathbb{R} : a \leq u \leq< b\}$ is an object very well known to all of us: it is none other than the closed interval $[a, b]$. So we can say the same thing as (11.196) by writing "$[a, b]$" instead of "$\{u \in \mathbb{R} : a \leq u \leq< b\}$". And we get

$$(\exists a \in \mathbb{R})(\exists b \in \mathbb{R})\Big([a, b] \neq \emptyset \wedge \int_a^b x^2 dx = c\Big), \tag{11.197}$$

which says exactly the same thing as (11.196). but now there is no "u" anymore.

(b) The definite integral $\int_a^b x^2 dx$ is a number that is completely determined by a and b. We do not need to ask "who is x?" in order to determine this number. Actually, the integral can be computed, and the result is $\frac{1}{3}(b^3 - a^3)$. So we can say the same thing as (11.197) by writing "$\frac{1}{3}(b^3 - a^3)$" instead of "$\int_a^b x^2 dx$", and we get

$$(\exists a \in \mathbb{R})(\exists b \in \mathbb{R})\Big([a, b] \neq \emptyset \wedge \frac{1}{3}(b^3 - a^3) = c\Big), \tag{11.198}$$

or, more nicely,

$$(\exists a \in \mathbb{R})(\exists b \in \mathbb{R})\Big([a, b] \neq \emptyset \wedge b^3 - a^3 = 3c\Big), \tag{11.199}$$

which say exactly the same thing as (11.197). but now there is no "x" anymore.

2. A variable is dummy if it can be replaced by any other variable (except with variables that are already being used for something else) without changing the meaning of the sentence. For example,

(a) If instead of the expression "$\{u \in \mathbb{R} : a \leq u \leq b\}$" we use a different letter and write something like "$\{v \in \mathbb{R} : a \leq v \leq b\}$", or "$\{z \in \mathbb{R} : a \leq z \leq b\}$", or maybe "$\{\alpha \in \mathbb{R} : a \leq \alpha \leq b\}$", or "$\{\diamond \in \mathbb{R} : a \leq \diamond \leq b\}$", nothing changes. So, for example, we can rewrite (11.196) as

$$(\exists a \in \mathbb{R})(\exists b \in \mathbb{R})\Big(\{q \in \mathbb{R} : a \leq q \leq b\} \neq \emptyset \wedge \int_a^b x^2 dx = c\Big), \tag{11.200}$$

which says exactly the same thing as (11.196). but now there is no u anymore.

(b) If we replace the definite integral $\int_a^b x^2 dx$ by the expression $\int_a^b h^2 dh$, or $\int_a^b \sigma^2 d\sigma$, or $\int_a^b m^2 dm$, nothing changes. So, for example, we can rewrite (11.200) as

$$(\exists a \in \mathbb{R})(\exists b \in \mathbb{R})\Big(\{q \in \mathbb{R} : a \leq q \leq b\} \neq \emptyset \wedge \int_a^b k^2 dk = c\Big), \quad (11.201)$$

which says exactly the same thing as (11.196). but now there is no u and no x anymore.

Summarizing: Sentence (11.196) defines a ***one-variable predicate, with the open variable*** c. So we can call this predicate $P(c)$.

And then we may ask: can we tell what this predicate $P(c)$ is? Can we find a simpler expression for $P(c)$?

It turns out that, in this case, the answer is "yes, we can":

$\boxed{P(c) \textbf{ \textit{just says} } \textit{“}c \geq 0\textit{”}}$.

Proof. We want to prove that $(\forall c \in \mathbb{R})(P(c) \iff c \geq 0)$.

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

We want to prove that $P(c) \iff c \geq 0$.

For that purpose, we will prove the implications $P(c) \implies c \geq 0$ and $c \geq 0 \implies P(c)$.

Proof that $P(c) \implies c \geq 0$.

Assume $P(c)$.

This means that

$$(\exists a \in \mathbb{R})(\exists b \in \mathbb{R})\Big([a, b] \neq \emptyset \wedge b^3 - a^3 = 3c\Big).$$

Pick real numbers a, b such that $a, b] \neq \emptyset$ and $b^3 - a^3 = 3c$.

Since $a, b] \neq \emptyset$, it follows that $a \leq b$. (Reason: if $a > b$ then the set $[a, b]$, i.e., the set $\{u \in \mathbb{R} : a \leq u \leq b\}$, would be empty.)

Since $a \leq b$, we have $a^3 \leq b^3$.

So $b^3 - a^3 \geq 9$.

So $3c \geq 0$.

Hence $\boxed{c \geq 0}$.

Proof that $c \geq 0 \implies P(c)$.

Assume that $c \geq 0$.

Let $a = 0$, $b = \sqrt[3]{3c}$.

Then $b \geq 0$.

So the closed interval $[a, b]$ (i.e., the interval $[0, b]$) is nonemtpy.

And $b^3 - a^3 = 3c$.

Hence $[a, b] \neq \emptyset \wedge b^3 - a^3 = 3c$.

So

$$(\exists a \in \mathbb{R})(\exists b \in \mathbb{R})\Big([a, b] \neq \emptyset \wedge b^3 - a^3 = 3c\Big).$$

That is, $\boxed{P(c) \text{ holds.}}$

Since we gave proved that $P(c) \implies c \geq 0$ and that $c \geq 0 \implies P(c)$, we can conclude that $P(c) \iff c \geq 0$.

Since we have proved that $P(c) \iff c \geq 0$ for arbitrary real c, we have proved that $(\forall c \in \mathbb{R})\Big(P(c) \iff c \geq 0\Big)$. **Q.E.D.**

11.3 First-order predicate calculus

The language we use in mathematics is a ***predicate calculus*** because it enables us to predicates. And it is ***first-order*** because we can quantify variables, and write things such as "$(\forall x \in P)x$ likes Mark" (meaning, if P is the set of all people, "everybody likes Mark"), but we cannot quantify over predicates. That is,

- We cannot say things such as "'for every predicate $P(x)$ and every predicate $Q(x)$ if $(\forall x)P(x)$ is true and $(\forall x)Q(x))$ is true, then if $(\forall x)(P(x) \wedge Q(x))$ is true."
- We can say this for a particular pair of predicates $P(x)$, $Q(x)$ (for example, we can say "if everybody likes coffee and everybody likes milk then everybdoy likes coffee and milk", or we can say "if everybody studies and everybody reads books then everybdoy studies and reads books"), but we cannot say the same thing for arbitrary predicates $P(x)$, $Q(x)$.

It turns out that there are "second order" languages, in which you can say things like "'for every predicate $P(x)$ and every predicate $Q(x)$ if $(\forall x)P(x)$ is true and $(\forall x)Q(x))$ is true, then if $(\forall x)(P(x) \wedge Q(x))$ is true." But the language we are using here is a ***first-order language***, in which those things cannot be said.

11.4 Logical connectives

In firts-order predicate calculus, one or more sentences can be combined to form other sentences. The symbols used to combine sentences are called the ***logical connectives.*** And there are exactly seven of them

11.4.1 The seven logical connectives

And here they are, in all their glory:

The seven logical connectives

1. The ***negation symbol*** $\sim$

   (meaning "no", "it's not the case that").
2. The ***conjunction symbol***, $\wedge$
   (meaning "and").
3. The ***disjunction symbol***, $\vee$
   (meaning "or").
4. The ***implication symbol***, $\Longrightarrow$
   (meaning "implies", or "if ... then").
5. The ***biconditional symbol***, $\Longleftrightarrow$
   (meaning "if and only if").
6. The ***existential quantifier symbol***, $\exists$
   (meaning "there exists ... such that", or "it is possible to pick ... such that").
7. The ***universal quantifier symbol***, $\forall$
   (meaning "for all",or "for avery", or "for an arbitrary").

11.4.2 How the seven logical connectives are used to form sentences

These seven symbols are used to form new sentences as follows:

1. The negation symbol $\sim$ is a ***one-argument connective***: it can be

put in front of a sentence A to form the sentence $\sim A$ (meaning "no A", or "it's not the case that A"). For example: "$\sim 3|5$" means "3 does not divide 5".

2. The conjunction symbol $\wedge$ is a ***binary connective***, or ***two-argument connective***: it can be put between two sentences A, B to form the sentence $A \wedge B$, (meaning "A and B"). For example: "$(\sim 3|5) \wedge 3|6$" means "3 does not divide 5 and 3 divides 6".

3. The disjunction symbol $\wedge$ is a ***binary connective***, or ***two-argument connective***: it can be put between two sentences A, B to form the sentence $A \vee B$, (meaning "A or B"). For example: "$x > 0 \vee x < 0$" means "$x > 0$ or $x < 0$".

4. The implication symbol $\Longrightarrow$ is a ***binary connective***, or ***two-argument connective***: it can be put between two sentences A, B to form the sentence $A \Longrightarrow B$, (meaning "A implies B", or "if A then B"). For example: "$x \neq 0 \Longrightarrow x^2 > 0$" means "if $x > 0$ then $x^2 > 0$".

5. The biconditional symbol $\Longleftrightarrow$ is a ***two-argument connective***, that is ***binary connective***: it can be put between two sentences A, B to form the sentence $A \Longleftrightarrow B$, (meaning "A if and only if B"). For example: "$(2|n \wedge 3|n) \Longleftrightarrow 6|n$" means "2 divides n and 3 divides n if and only if 6 divides n".

6. The existential quantifier symbol $\exists$ has a more complicated grammar:

   (a) Using $\exists$ we can form ***existential quantifiers***.

   (b) There are two kinds of existential quantifiers:

      i. ***Unrestricted existential quantifiers*** are expressions
      $$\Big(\exists x\Big),$$
      that is: left parenthesis, $\exists$, variable, right parenthesis.

      ii. ***Restricted existential quantifiers*** are expressions
      $$\Big(\exists x \in S\Big),$$
      that is: left parenthesis, $\exists$, variable, $\in$, name of a set, right parenthesis.

   (c) Then we can take a sentence A (or $A(x)$) and put a restricted or unrestricted existential quantifier in front, forming the sentences $(\exists x)A$ ("there exists x such that A", or "it is possible to pick

x such that A") and $(\exists x \in S)A$ ("there exists x belonging to S such that A", or "it is possible to pick x belonging to S such that A").

7. The universal quantifier symbol $\forall$ has a grammar similar to that of the existential quantifier symbol:
   (a) Using $\forall$ we can form ***universal quantifiers***.
   (b) There are two kinds of universal quantifiers:
      i. ***Unrestricted universal quantifiers*** are expressions
      $$\big(\forall x\big),$$
      that is: left parenthesis, $\forall$, variable, right parenthesis.
      ii. ***Restricted universal quantifiers*** are expressions
      $$\big(\forall x \in S\big),$$
      that is: left parenthesis, $\forall$, variable, $\in$, name of a set, right parenthesis.
   (c) Then we can take a sentence A (or $A(x)$) and put a restricted or unrestricted universal quantifier in front, forming the sentences $(\forall x)A$ ("for all x, A", or "A i strue for arbitrary x") and $(\forall x \in S)A$ ("for all x belonging to S, A", or "A is true for arbitrary x in S").

11.5 Conjunctions ("$\wedge$", i.e., "and")

The symbol

$$\bigwedge$$

is the ***conjunction symbol***, and means "and".

Hence,

- 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***.
- In a conjunction $P \wedge Q$, the sentences P, Q are the ***conjuncts***.

11.5.1 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.** □

11.5.2 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 207, where we talked about another "stupid rule", namely, Rule $\wedge_{prove}$. That remark also applies to Rule $\wedge_{use}$. □

11.6 Disjunctions ("$\vee$", i.e., "or")

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***.
- In a disjunction $P \vee Q$, the sentences P, Q are the ***disjuncts***.

11.6.1 The meaning of "or" in mathematics

The meaning of "or" in mathematics

In English, when we use the word "or", it can have two different meanings:

1. ***Inclusive*** "or", that is, "one or the other or both".

or

2. ***Exclusive*** "or", that is, "one or the other but not both".

For example, if a store announces that

> If you are a student or a senior citizen then you are entitled to a 15% discount on your purchases.

then, obviously, anyone who is both a student and a senior citizen will be entitled to a discount. So this is an example of ***inclusive*** or.
On the other hand, if a restaurant waiter asks you "would you like tea or coffee?", then it is clear that you can have one or the other but not both, so this an example of ***exclusive*** or.

In mathematics, "or" is always inclusive.

So, if I say, for example,

> if a and b are integers and a is even or b is even, then the product ab is even,

then this also applies to the case when both a and b are even.

11.6.2 The truth table for "or"

We can summarize what we said about "or" in the previous section by means of the following "truth table" for the connective "$\vee$":

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

This truth table says that, if P, Q are propositions, then:

1. "$P \vee Q$" is false when P and Q are both false;

2. "$P \vee Q$" is true in all other cases, that is, when P and Q are both true, or when one of them is true and the other one is false.

11.6.3 Using a disjunction: the "proof by cases" rule

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 52. Suppose you want to prove that

$$(\forall x \in \mathbb{R})(x \neq 0 \Longrightarrow x^2 > 0)\,. \tag{11.202}$$

Then you could reason as follows. Since $x \neq 0$, there are two possibilities: $0 < x$ or $x < 0$. So

$$0 < x \ \vee \ x < 0\,. \tag{11.203}$$

Sincde we have the disjunction (11.203), we are in a position to use Rule $\vee_{use}$. To do this, we consider each of the two possibilities "$0 < x$" and "$x < 0$" separately.

> First we assume that $\boxed{0 < x.}$
>
> Then we use the fact that we can multiply both sides of an inequality by a positive number[64]. Since $0 < x$ (because we are assuming that $0 < x$), we can multiply both sides of "$0 < x$" by x, and get $x.0 < x.x$.
>
> But $x \cdot 0 = 0$ by a well-known theorem[65]
>
> And $x{\cdot}x = x^2$. (This is because the definition of x^2 says that $x^2 = x{\cdot}x$.)
>
> So $\boxed{0 < x^2}$.
>
> Next we assume that $\boxed{x < 0.}$
>
> Then we use the axiom that says that we can add a real number to both sides of an inequality and the result is an inequality going in the same direction[66]. So we add $-x$ to both sides of "$x < 0$" and get $0 < -x$.

[64] This is one of the axioms of real number theory, that we will discuss later. The axiom says: $(\forall x \in \mathbb{R})(\forall y \in \mathbb{R})(\forall z \in \mathbb{R})\Big((x < y \wedge 0 < z) \Longrightarrow xz < yz\Big)$.

[65] The theorem says that $(\forall x \in \mathbb{R})x.0 = 0$.

[66] Precisely, the axiom says: $(\forall x \in \mathbb{R})(\forall y \in \mathbb{R})(\forall z \in \mathbb{R})(x < y \Longrightarrow x + z < y + z\Big)$.

Then we use the axiom about multiplication of 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 $x \cdot 0 = 0$.

And $(-x) \cdot (-x) = x \cdot x$.

So $0 < x \cdot x$.

And $x \cdot x = x^2$, by the definition of "square".

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 $\boxed{\boxed{0 < x^2}}$.

What we have done in this example is this: we knew that a disjunction $A \vee B$ was true. (In our example, A was "$0 < x$" and B was "$x < 0$".) Then we proved that a ceartain conclusion C must hold if A is true, and also if B is true. (In our example, C was "$0 < x^2$".) Then we concluded that C must be true. And the reason is quite simple: one of A, B is true, and in either case C is true, so C is true.

This is exactly what the proof by cases rule says.

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.

11.6.4 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 $P \vee Q$.

Example 53. Let us prove that

$$(\forall n \in \mathbb{Z})(3|n \vee 3|n^2 - 1)\,. \tag{11.204}$$

Proof.

Let n be an arbitrary integer.

We want to prove that $3|n \vee 3|n^2 - 1$.

Assume that $\sim 3|n$, that is, 3 does not divide n.

We want to prove that $3|n^2 - 1$.

Clearly, $n^2 - 1 = (n-1)(n+1)$.

Furthermore, it is well known that if k, $k+1$ and $k+2$ are any three consecutive integers, then one of them must be divisible by 3.

Applying this with $k = n - 1$, we see that one of the integers $n-1, n, n+1$ is divisible by 3.

But we are assuming that n is not divisible by 3.

Hence one of the numbers $n-1, n+1$ is divisible by 3.

So the product $(n-1)(n+1)$ is divisible by 3.

That is, $n^2 - 1$ is divisible by 3.

So we have proved that $3|n^2-1$, assuming that $\sim 3|n$.

By Rule $\vee_{prove}$, it follows that $3|n \vee 3|n^2-1$.

We have proved that $3|n \vee 3|n^2-1$ for an arbitrary integer n.

Therefore $(\forall n \in \mathbb{Z})3|n \vee 3|n^2-1$. **Q.E.D.**

11.7 Implications ("$\Longrightarrow$", i.e., "if ... then")

Implication: The symbol

$$\Longrightarrow$$

is the ***implication symbol***, and means "implies".

A sentence "$P \Longrightarrow Q$" is read as

$$P \text{ implies } Q$$

or as

$$\text{If } P \text{ then } Q\,.$$

Then

- 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***.

- In a conditional sentence $P \Longrightarrow Q$, P is the ***premiss*** (or ***antecedent***), and Q is the ***conclusion*** (or ***consequent***.

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

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**[67]

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

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

Example 54. Suppose you know that

P1. If you are a student then you are entitled to a discount,

and you also know that

P2. you are a student.

Then you can conclude that

C. you are entitled to a discount. □

11.7.2 The "for all...implies" combination

One of the most important and widely used combinations of moves in proofs is what we may call ***the "for all...implies" combination***.

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

$$(\forall x \in S)(A(x) \Longrightarrow B(x))\,. \tag{11.205}$$

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

$$A(a)\,. \tag{11.206}$$

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

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

  Step 1: Use the specialization rule to go from (11.205) to

  $$A(a) \Longrightarrow B(a)\,. \tag{11.207}$$

  Step 2: Use Modus Ponens to go from (11.207) and (11.206) to

  $$B(a)\,. \tag{11.208}$$

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

$$(\forall x)(A(x) \Longrightarrow B(x))\,. \tag{11.209}$$

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" combination 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 (11.209), 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" combination, 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" combination, 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]

11.7.3 Proving an implication (Rule $\Longrightarrow_{prove}$)

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 55. 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. □

11.7.4 The connectives "$\wedge$" and "$\Longrightarrow$" are very different

Students sometimes think that "If P then Q" is basically the same as "P and Q", or "P then Q". But this is very wrong and it important that you should understand the difference between "P and Q" and "If P then Q".

Take, for example, the sentences

Today is Friday and tomorrow is June 12.

and

If today is Friday then tomorrow is June 12.

Using "P" to represent the sentence "Today is Friday" and "Q" to represent the sentence "Tomorrow is June 2", the first sentence is $P \wedge Q$, and the second one is $P \Longrightarrow Q$.

What conditions have to be satisfied for $P \wedge Q$ to be true? What conditions have to be satisfied for $P \Longrightarrow Q$ to be true?

The sentence $P \wedge Q$ is true if both P and Q are true. In our example, the only way the sentence "Today is Friday and tomorrow is June 12" can be true is if today is Friday and tomorrow is June 12, So ***the***

sentence "Today is Friday and tomorrow is June 12" is true if today is Friday June 11, and in no other case.

On the other hand, ***The sentence $P \Longrightarrow Q$ when Q is true, and also when P is false. And if neither one of these conditions hold (that is, if Q is false and P is true) then $P \Longrightarrow Q$ is false.*** So, in our example, the only possible situation when "If today is Friday then tomorrow is June 12" would be false is if today is Friday but tomorrow is not June 12. So ***the sentence "If today is Friday then tomorrow is June 12" is true if today is not Friday, is also true if tomorrow is June 12, and is false if today is Friday but tomorrow is not June 12.***

We can summarize these observations by means of the following "truth tables" for the connectives "$\wedge$" and "$\Longrightarrow$":

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

P	Q	$P \Longrightarrow Q$
T	T	T
T	F	F
F	T	T
F	F	T

The first table gives you the truth value[68] of $P \wedge Q$ in terms of the truth values of P and Q, and the second table gives you the truth value of $P \Longrightarrow Q$ in terms of the truth values of P and Q.

Notice that ***what makes the truth tables for "$wedge$" and "$\Longrightarrow$" is the last two lines.*** In particuler:

> ***$P \Longrightarrow Q$ is always true when Q is true, no matter whether P is true or false.***

and

[68] Every sentence, when used correctly, has a ***truth value***: the truth value is T is the sentence is true, and F is the sentence is false. For example: the truth value of "$3 > 5$" is F, the truth value of "$3 < 5$" is T. How about the truth value of "$x < 5$". If you tell me that $x < 5$ without having told me who x is, then I do not knwo the truth value of "$x < 5$". But this would be an incorrect us of "$x < 5$". If you were writing a proof, then you could never have "$x < 5$" as one of the steps, unless you have told the reader before, in some previous step, who x is, and once you have done that, the truth value of "$x < 5$" would be known. For example, if you said in a previous step "Let $x = \frac{1+\sqrt{5}}{2}$", then I would know that "$x < 5$" is true. (Proof: $\sqrt{5} < 5$. So $1 + \sqrt{5} < 6$. So $\frac{1+\sqrt{5}}{2} < 3$. Hence $\frac{1+\sqrt{5}}{2} < 5$. So $x < 5$.)

$P \Longrightarrow Q$ is always true when P is false, no matter whether Q is true or false.

So for example, the following sentences are true:

- If the Earth is a planet then 3 is a prime number.
- If the Earth is a comet then 3 is a prime number.
- If the Earth is a comet then 6 is a prime number.

The first one and the second one are true because the conclusion (that is, "3 is a prime number) is true. . (It does not matter, for the second sentence, that the premiss—"the Earth is a comet"— is false.)

And the second one and third one are true because the premiss ("the Earth in a comet" is false. (It does not matter whether for the second sentence, that the conclusion—"6 is a prime number"— is false.)

On the other hand, the sentence "If the Earth is a planet then 6 is a prime number" is false, because the premiss ("The Earth is a planet") is true, but the conclusion ("6 is a prime number") is false.

11.7.5 Isn't the truth table for $\Longrightarrow$ counterintuitive?

Students often ask questions about the implication connective $\Longrightarrow Q$ and in partuclar about the truth table for the implication.

One often raise question is "how can '$P \Longrightarrow Q$' be true if P and Q have nothing to do with each other?".

For example, we said that the sentence "If the Earth is a planet then 3 is a prime number" is true, but what does the fact that the Earth is a planet have to do with 3 being a prime number? That sounds like a good question, but let us think about it. I suggest that you do do this:

Think of "$P \Longrightarrow Q$" as saying "it does not happen that P is true without Q also being true".

In other words: what "$P \Longrightarrow Q$" does is exclude the possibility that you might ever run into a "bad situation", menaing, "a situation where P is true

but Q is not". And this is the only possibilty excluded the implication. So, in particular,

- if P is false then you will not be in a bad situation, so "$P \Longrightarrow Q$" is true.
- if Q is true then you will not be in a bad situation, so "$P \Longrightarrow Q$" is true.

Once you understand this, you will see that it does not matter very much whether P and Q have something to do with each other. Maybe $P$and Q are totally unrelated, but if, for example, they both happen to be true then "$P \Longrightarrow Q$" is true. And also, "$P \Longrightarrow Q$" will be true if both P and Q are false, or if P is false and Q is true.

Example 56. Suppose a street sign says:

IF YOU ARE DRIVING AT MORE THAN 25MPH YOU WILL GET A FINE.

Supoose you want to prove to a friend of yours that the municipal government that put up the sign isn't really enforcing its own rule. What do you have to do to prove this?

Let "P" represent the premiss, i.e., "you are driving at more than 25mph", and let "Q" represent the conclusion, that is, "you will get a fine". Then the street sign asserts the implication "$P \Longrightarrow Q$".

Certainly,

- If you find someone driving at 20mph, that will do nothing to prove your case. ***That's because in that case the implication "$P \Longrightarrow Q$" is true, according to the truth table for the implication.*** It does not matter whether that driver got a fine or not[69].
- If you find someone who got a fine, that will do nothing to prove your case. ***That's because in that case the implication "$P \Longrightarrow Q$" is true, according to the truth table for the implication.*** It

[69]The driver may have been given a fine for some other reason, e.g., using a cell phone while driving.

does not matter whether that driver was driving at more than 25mph or not.[70].

- The only way to prove that the injunction in the street sign is not being enforced is to find cases of drivers that were driving at more than 25mph but did not get a fine. ***That's because the onlt case when the implication "$P \implies Q$" is false, according to the truth table for the implication, is when the premiss is true but the conclusion is false.***

Example 57. Alice is a cashier at a department store, and she has to follow the rule that

IF A CUSTOMER PAYS CASH FOR A PURCHASE THEN ALICE HAS TO PUT THE MONEY SHE COLLECTED IN A DRAWER.

Suppose you are a detective and you want to prove that Alice is not obeying the rule. What do you have to do?

- If you find a situation when there was not customer at all, or there was customer that did not pay cash, then that will do nothing prove your case. ***That's because in that case the implication "$P \implies Q$" is true, according to the truth table for the implication.*** It does not matter whether Alice put money is the drawer or not[71].

- If you find a situation where Alice put cash in the drawer even though she did not collect any money from a customer, then that will do nothing to prove your case. ***That's because in that case the implication "$P \implies Q$" is true, according to the truth table for the implication.*** It does not matter that there was no customer poaying cash[72].

[70] The driver may have been driving at 20mph but may have been given a fine for some other reason, e.g., using a cell phone while driving.

[71] Why would Alice have put money in the drawer if she did not collect any cash from the customer? Who knows?

[72] Again, why would Alice put money in the drawer even if she did not collect the money from a customer? Who knows? And who cares? The point is: ***even if she put money in the drawer when there had been no customer that paid her the money, so P was false but Q was true, she did not violate the rules.***

- The only way you can prove that Alice is violating the rules is by showing that a customer paid cash but Alice did notput themoney in the drawer. ***That's because the only case when the implication "$P \implies Q$" is false, according to the truth table for the implication, is when the premiss is true but the conclusion is false.***

Example 58. Suppose you have a natural number n, but you do not know which number it is. (For example, maybe someone gave you a sealed envelope containing a card where the number is written. So the number is there, it's a fixed number, but you just do not knwo which specific number it is.)

Suppose you are asked to prove that

(*) If n is even then n^2 is divisible by 4.

Then you could ask: could (*) possibly be false? Could there be a possible value of n for which (*) is false. (Remember that you do not know who n is. So if you want be able to assert for sure that (*) is true you have to consider all possible values of n. If you find one value of n for which (*) is not true, then you cannot be sure that n is true, because the number that you have in the envelope could be the one you have found, the one for which (*) is false. But if you can make sure that no such number exists, then you can be sure that (*) is true, even though you do not know who n is.)

What would have to happen for (*) to be false? Well, according to our truth table, the only case when the implication (*) is false is when the premiss is true but the conclusion is not. So to make sure that (*) is true, you have to consider numbers n that are even, because if n is not even then (*) is true. You indicate that you are going to do that by writing:

Assume that n is even.

(In other words: ***you are allowed to assume that n is even because if n is not even then (*) is automatically true thanks to the truth table for the implication.***)

And then you move on to prove that n^2 is divisible by 4. (Since n is even, we can pick a natural number k such that $n = 2k$. Then $b^2 = 4k^2$, so n^2 is divisible by 4.)

And now you can be sure that (*) is true. The number n is even or odd, but in either case (*) is true, even though in each case it's true for a different reason: if n is not even, then (*) is true because of the truth table for the implication, and if n is even then (*) ia true because in that case we have proved that the conclusion (that is, "n^2 is divisible by 4") must be true.

Finally, we have prove that (*) must be true for any natural number, because we have proved for n, but n could be any number. So we can conclude that

$$(\forall n \in \mathbb{N})\Big(n \text{ is even} \Longrightarrow n^2 \text{ is divisible by } 4\Big)\,,$$

or, if you prefer,

$$(\forall n \in \mathbb{N})\Big(2|n \Longrightarrow 4|n^2\Big)\,.$$

So we can structure our proof as follows:

THEOREM. $(\forall n \in \mathbb{N})\Big(2|n \Longrightarrow 4|n^2\Big)\,.$

PROOF We want to prove that $(\forall n \in \mathbb{N})\Big(2|n \Longrightarrow 4|n^2\Big)\,.$

Let $n \in \mathbb{N}$ be arbitrary.

We want to prove that $2|n \Longrightarrow 4|n^2$.

Assume that $2|n$.
Then $(\exists k \in \mathbb{N})n = 2k$.
Pick one such k and call it k_*.
Then $k_* \in \mathbb{N}$ and $n = 2k_*$.
Then $n^2 = (2k_*) \cdot (2k_*) = 4k_*^2$.
Let $q = k_*^2$.
Then $n^2 = 4q$.
So $(\exists k)n^2 = 4k$.
Hence $4|n^2$.

We have proved that $4|n^2$ assuming that $2|n$. Hence

$$2|n \Longrightarrow 4|n^2\,.$$

We have proved that $2|n \Longrightarrow 4|n^2$ for an arbitrary n. Therefore

$$(\forall n \in \mathbb{N})\Big(2|n \Longrightarrow 4|n^2\Big)\,.$$

Q.E.D.

I hope that these remarks will suffice to clarify they way implication works. Implication will be discussed in great detail later.

11.8 Biconditionals ("$\iff$", i.e., "if and only if")

The ***biconditional*** is the symbol

$$\iff .$$

It is a *binary connective*, like $\wedge$, $\vee$, and $\implies$. That means that $\iff$ ***can be used to connect two sentences***.

If P and Q are sentences, the sentence "$P \iff Q$" is read as

$$\boxed{P \text{ if and only if } Q}$$

or

$$\boxed{P \text{ is equivalent to } Q}\,.$$

And mathematicians often use "iff" as shorthand for "if and only if", so they write "P iff Q."

$$\boxed{P \text{ iff } Q}\,.$$

The precise meaning of "equivalence" will be explained later. But, if you want to know right away what it means, it's very simple:

When you know that P is equivalent to Q then you can pass freely from P to Q. That is, if you know that P is true then you can write Q, and if you know that Q is true then you can write P.
So for all practical purposes if "$P \iff Q$" is true then P and Q are interchangeable.

11.8.1 The meaning of "if and only if"

You should think of "P iff Q" as meaning

$$(P \iff Q) \wedge (Q \iff P)\,.$$

That is, "$P \iff Q$" means[73]

$$\text{If } P \text{ then } Q \text{ and if } Q \text{ then } P\,,$$

[73] *This note is only for philosophically minded nitpickers.* What does "means" mean? The point of view adopted here is that *the meaning of a word, phrase or symbol consists of the rules for the use of that word, phrase or symbol.* For example, the meaning of "and" is the specification that if P, Q are two sentences, then (i) if you have "P and Q" you can go to P and you can go to Q, and (ii) if you have P and you have Q then you can go to "P and Q." That is, *the meaning of "and" is captured by Rules $\wedge_{use}$ and $\wedge_{prove}$.* Naturally,

or

$$P \text{ implies } Q \text{ and } Q \text{ implies } P\,.$$

In order to make this true, we will choose the rules for proving and using biconditional sentences as follows:

- *To prove "$P \iff Q$" you do exactly the same thing that you would do to prove $(P \iff Q) \wedge (Q \iff P)$.*
- *To use "$P \iff Q$" you do exactly the same thing that you would do to use $(P \iff Q) \wedge (Q \iff P)$.*

So, for example, suppose you want to prove that

$$(\forall x \in \mathbb{R})\Big(x^2 = 4 \iff (x = 2 \vee x = -2)\Big)\,. \tag{11.210}$$

Then you would start by introducing into your proof an arbitrary real number called x, and then you would prove that

$$(x^2 = 4 \iff (x = 2 \vee x = -2)\,. \tag{11.211}$$

And to prove (11.211), which is an "iff" sentence, you would prove both implications $x^2 = 4 \implies (x = 2 \vee x = -2)$ and $(x = 2 \vee x = -2) \implies x^2 = 4$. (The proof of these two sentences is very simple: to prove that $x^2 = 4 \implies (x = 2 \vee x = -2)$, you use the fact that a positive real number r cannot have more than two square roots[74]. Since 2 and -2 are two distinct square roots

this does not cover all the uses of "and" in our culture, such as, for example, to indicate a progression (as in "this is getting better and better"), or to indicate a causal relation, (as in "do that and I'll hit you"), or the literary use full of nuances (as 'in 'tomorrow and tomorrow and tomorrow"). And, most importantly for us, it does not cover the use of "and" to connect *nouns*, as in "slings and arrows". But it's what "and" means in logic and mathematics. If you want to program a computer so that it will know what "and" means, you have to tell the computer how to use "and". And this amounts to programming the computer to use rules $\wedge_{use}$ and $\wedge_{prove}$. And you don't need to tell the computer anything else. A similar situation arises with the biconditional. A computer that "knows" the rules $\iff_{use}$ and $\iff_{prove}$ "knows" all it needs to know to work with the biconditional, and for that reason I believe that knowing the meaning of "$\iff$" amounts to knowing the two rules for working with it.

[74]This was proved in the notes for Lectures 2,3,4 but, just in case, here is a quick proof: suppose r has three distinct square roots a, b, c. Then $a^2 = r$, $b^2 = r$ and $c^2 = r$. Hence $a^2 - b^2 = 0$. So $(a - b)(a + b) = 0$. Therefore $a - b = 0$ or $a + b = 0$. Since a and b are different, it cannot be the case that $a - b = 0$, so $a + b$ must be zero, and then $b = -a$. Now we can use exactly the same argument with c instead of b, and conclude that $c = -a$. But then $c = b$, contradicting the fact that $b \neq c$.

of 4, there cannot be a third square root. So, if $x^2 = 4$, so x is a square root of 4, it follows that x must be 2 or -2. So $x^2 = 4 \Longrightarrow (x = 2 \vee x = -2)$. To prove the other implication, i.e., that $(x = 2 \vee x = -2) \Longrightarrow x^2 = 4$, just observe that if $x = 2$ then $x^2 = 4$, and if $x = -2$ then $x^2 = 4$ as well,)

11.8.2 The rules for proving and using biconditionals

Now let us state explicitly the rules for proving and using biconditional sentences.

As I explained in the previous subsection, *these rules are designed so as to make "$P \Longleftrightarrow Q$" mean precisely what we want it to mean, that is "$(P \Longrightarrow Q) \wedge (Q \Longrightarrow P)$".*

The rules are as follows.

Rule $\Longleftrightarrow$ *prove*

If P, Q are sentences, and you have proved the sentences

$$P \Longrightarrow Q$$

and

$$Q \Longrightarrow P\,,$$

then you can go to

$$P \Longleftrightarrow Q\,.$$

Rule $\Longleftrightarrow$ *use*

If P, Q are sentences, and you have proved the sentence

$$P \Longleftrightarrow Q\,,$$

then you can go to

$$P \Longrightarrow Q$$

and you can also go to

$$Q \Longrightarrow P\,.$$

11.9 The other six rules

So far I have given you eight rules, two for each of the connectives $\wedge$, $\vee$, $\Longrightarrow$, and $\Longleftrightarrow$.

In addition, there are six more rules that we have already discussed:

1. Rule $\forall_{prove}$, the rule for proving a universal sentence. (This rule is sometimes called "universal generalization".)
2. Rule $\forall_{use}$, the rule for using a universal sentence. (This is sometimes called the "specialization rule".)
3. Rule $\exists_{prove}$, the rule for proving an existential sentence.. (This rule is sometimes called the "existential generalization rule".)
4. Rule $\exists_{use}$, the rule for using a universal sentence. (This rule is sometimes called the "existential specialization rule".)
5. The proof by contradiction rule.
6. Rule SEE, substitution of equals for equals (also called "Rule $=_{use}$").

So ***we now have all fourteen rules!***

11.10 Are the logical rules hard to understand and to learn and remember ?

Most of the logical rules are very simple and easy to remember. For example,

- The rules for using and proving $\wedge$ sentences are so stupid that you might object to having them because they are so obvious, but you certainly cannot find it hard to understand them.
- The rules for using and proving universal sentences are also natural:
  - if you know that all the items in this store cost 1 dollar, and you pick an item in this store, you can be sure that it costs 1 dollar. That's all that Rule $\forall_{use}$ says.
  - if you prove that a schmoo must be green, without using any information about that schmoo other than the fact that it is a schmoo, then you can conlude that all schmoos are green. And that's= all that Rule $\forall_{prove}$ says.
- And the rules for using and proving existential sentences are natural as well:
  - if you know that somewhere in this store there is a schmoo, then you can go and get a schmoo and call it any way you want, for example "my woderful schmoo". That's all that Rule $\exists_{use}$ says.
  - if you find a schmoo, then you can conclude that schmoos exist. And that's all that Rule $\exists_{prove}$ sats.

11.10.1 Proofwriting and rules for proofs

Writing proofs is like playing chess, checkers, or some other board game.

- There are rules that tell you which moves are allowed. (Notice that the rules for proofs never say "you ***have*** to do this". They say "you ***are allowed*** to do this". It's exactly like the moves you are allowed to make in a board game.)
- You have to obey the rules all the time.
- If you cheat, by violating the rules once, then you are out of the game.
- If you know how to play, you will never make a move that violates the rules.
- Once you know the moves, then the hard part begins: you have to figure out how to choose which moves to make in order to win. And that is where proofwriting becomes difficult and challenging: some people are better than others at figuring out how to win.
- From 1637 until 1995, many mathematicians tried very hard to prove Fermat's last theorem. Finally, Andrew Wiles suceeded in doing it in 1995.
- But the proofs we do in this course are not that hard.

12 Induction

12.1 Introduction to the Principle of Mathematical Induction

You know that the following is true:

(*) *Every integer is even or odd, and not both.*

How can we prove statement (*) ?

First, we have to make it clear what we mean by "even" and "odd".

Definition 17.

1. *An integer n is <u>even</u> if n is divisible by 2, that is, if there exists an integer k such that $n = 2k$.*
2. *An integer n is <u>odd</u> if $n - 1$ is even, that is, if there exists an integer k such that $n = 2k + 1$.* □

Now that we know what it means for an integer to be "even" or "odd", we can try to prove some facts about even and odd integers. Here are some simple examples of theorems about even and odd numbers that are easy to prove:

Theorem 17. *If m and n are even integers, then $m + n$ is even. (That is, "the sum of two even integers is even".)*

Theorem 18. *If m and n are odd integers, then $m + n$ is even. (That is, "the sum of two odd integers is even".)*

Theorem 19. *If m and n are integers, m is even and n is odd, then $m + n$ is odd. (That is, "the sum of an even integer and an odd integer is odd".)*

Theorem 20. *If m and n are integers, and m or n is even, then $m.n$ is even. (That is, "the product of an even integer and an integer is an even integer".)*

Theorem 21. *If m and n are odd integers, then $m.n$ is odd. (That is, "the product of two odd integers is odd".)*

Theorem 22. *The integer 1 is odd and is not even.*

Theorem 23. *If an integer n is even, then the integers $n+1$ and $n-1$ are odd.*

Theorem 24. *If an integer n is odd, then the integers $n+1$ and $n-1$ are even.*

All these theorems are very easy to prove. I will do two of the proofs, and I will ask you to do all the others.

Proof of Theorem 18:

Let m, n be integers.

Assume m and n are odd.

We want to prove that $m+n$ is even.

Since m is odd, we can pick an integer j such that $m = 2j+1$.

Since n is odd, we can pick an integer k such that $n = 2k+1$.

Then $m+n = (2j+1)+(2k+1)$, so $m+n = 2j+2k+2$ and then $m+n = 2(j+k+1)$.

Hence $(\exists i \in \mathbb{Z})m+n = 2i$.

So $m+n$ is even. **Q**.**E**.**D**.

Proof of Theorem 22: First, we show that 0 is even. To prove this, we observe that $0 = 2.0$, so $(\exists k \in \mathbb{Z})0 = 2k$, and then 0 is even.

It then follows immediately that 1 is odd, because the definition of "odd integer" says that "n is odd" means "$n-1$ is even", so in particular "1 is odd" means "$1-1$ is even", and this is true, because $1-1=0$, and 0 is even.

Finally, we have to show that 1 is not even. For this purpose, we have to show that there is no integer k such that $2k = 1$. But there is only one real number k such that $2k = 1$, and that number is $\frac{1}{2}$, which is not an integer. So there is no integer k such that $2k = 1$. Hence 1 is not even. **Q**.**E**.**D**.

Problem 56. ***Prove*** Theorems 17, 19, 20, 21, 23, and 24.

WARNING: We have not proved yet that "odd" is equivalent to "not even". This will be proved later, in Theorem 29 in Section 12.3.3. But ***until we have proved it we cannot use it.*** So, for example, you are ***not*** allowed to prove that an integer n is even by contradiction, by saying "suppose n is not even, then n is odd." You cannot do that because we have not proved yet that "n is not even" is equivalent to "n is odd". □

What we actually want is to prove (*), i.e., to show that every integer is even or odd and not both.

Let us call an integer "good" if it is even or odd and not both even and odd. So we want to prove that

(**) *Every integer is good.*

We are going to prove first that every natural number is good, and then we will take the extra step of proving that every natual number is good.

So let us start by trying to prove that every natural number is good.

We already know that 1 is good. How about 2?

Theorem 25. *The number* 2 *is even and not odd. So* 2 *is good.*

Proof. 1 is odd, so by Theorem 24, $1 + 1$ is even, so 2 is even.

On the other hand, 2 cannot be odd, because if 2 was odd then $2 - 1$ would be even by Theorem 24.

So 2 is even and not odd, and then 2 is good. **Q**.**E**.**D**.

How about 3?

Theorem 26. *The number* 3 *is odd and not even. So* 3 *is good.*

Proof. 2 is even. So by Theorem 23, $2 + 1$ is odd, so 3 is odd.

On the other hand, 3 cannot be even, because if 3 was even then $3 - 1$ would be odd by Theorem 23, i.e., 2 would be odd.

So 3 is odd and not even, and then 3 is good. **Q**.**E**.**D**.

It is clear that we could go on the same way, and prove that 4 is good, 5 is good, 6 is good, *and so on.* And then we would conclude that every natural number is good.

However, saying "and so on" is not a rigorous way to ***prove*** that every natural number is good.

The key idea is this: we are going to prove that ***goodness is a property that is passed on from each natural number*** n ***to the number following it, i.e.,*** $n + 1$.

Precisely, we are going to prove:

Theorem 27. *If* n *is natural number and* n *is good, then* $n + 1$ *is good.*

Once we have proved Theorem 27, since we have already proved Theorem 22, which says that 1 is good, we will be able to reason as follows:
We know that

1. 1 ***is good***.
2. ***Goodness is passed on from each natural number*** n ***to its successor*** $n+1$**.** (That is: if $n \in \mathbb{N}$ and n is good, then $n+1$ is good.)

Then:

1. 2 is good, because 1 is good and 1 passes on the goodness property to 2,
2. 3 is good, because 2 is good and 2 passes on the goodness property to 3,
3. 4 is good, because 3 is good and 3 passes on the goodness property to 4,
4. 5 is good, because 4 is good and 4 passes on the goodness property to 5,

   ...

   and so on,

   so every natural number is good.

But it would be much better not to rely on vague phrases like "and so on", and to have instead a precise, rigorous way of doing the proof.

The key point is that ***all the natural numbers are eventually arrived at by counting***, so that, if we know that something is true for $n = 1$, and when we count (that is, go from 1 to 2, then from 2 to 3, then from 3 to 4, "and so on", each time passing from a natural number n to its successor $n + 1$), then at each step the goodness property will be passed on from n to $n + 1$, and eventually every natural number n will be reached by our counting process, so n will be good.

This means that

> Every property that is true of the number 1 and is passed on from each natural number to its successor must be true of all natural numbers.

And ***this is exactly what the Principle of Mathematical Induction (PMI) says***.

Example 59. Suppose you decide to paint natural numbers green according to the following rule: first, you paint the number 1 green. And then every time you paint a number n green, you go to its successor $n + 1$ and paint it green. Then the PMI guarantees that every natural number is painted green. □

Example 60. Suppose there is an infinitely long queue of people standing in line: person No. 1, then person No. 2, then person No. 3, then person

No. 4, and so on[75]. Suppose you have a flyer with an announcement that you want all the people in the queue to read. (For example, a message saying something like "if you come to my restaurant after the show you will get a great meal with a 20% discount"). Suppose you want everybody to read the flyer, but you have only one copy. Then all you have to do is

(1) Give the flyer to person No. 1,

and

(2) Make sure that each person passes on the flyer to the person next in line after reading it[76].

The PMI says the obvious thing: if you do (1) and (2) then everybody will eventually get your flyer. □

12.2 The Principle of Mathematical Induction (PMI)

As explained in the previous section, the ***Principle of Mathematical Induction (PMI)*** captures as a precise mathematical statement the intuitively clear fact that when we count ***we get all the natural numbers***.

Remark 14. There are other numbers (that is, people have invented other numbers), such as zero, the negative numbers -1, -2, etc., fractions such as $\frac{2}{3}$, $\frac{22}{7}$,$-\frac{5}{2}$, 2.75, -5.16, and even "irrational numbers", that cannot be expressed as fractions. But ***we do not get these numbers by the counting process***.

So, if you prove by induction that a statement $P(n)$ is true for all natural numbers, then it does ***not*** follow that it will be true for $n = 0$, because 0 is not a natural number, so if you count $1, 2, 3, 4, \ldots$ you will never get to 0.

[75]Sure, I am talking about an infinitely long queue, with infinitely many people. And you may object that this is impossible in reality. I have two answers to that. ANSWER NO. 1: This may be impossible in reality, but you can certainly ***imagine*** it! It may be impossible in reality for a person to jump 50 feet high, but you can certainly imagine Wonder Woman doing it, so why not imagine an infinite queue? ANSWER 2: Suppose you only have a finite queue, say 40 people. Then you can consider the following property $P(n)$ of a natural number: "person n got the message or there is no person n". This makes sense of every natural number n. If you guarantee that $P(n)$ is true of every natural number n, this will imply that persons 1, 2, 3, and so on up to person 40, will get the message. Property $P(n)$ will be true of every n but for different reasons: for $n = 1, 2, 3, 4, \cdots$, up to $n = 40$, it will be true because person n gets the message. And for larger n it will be true because there is no person No. n.

[76]For example, you could include in the flyer, in big letters, the statement PLEASE PASS THIS ON TO THE PERSON NEXT IN LINE TO YOU.

And it does not follow either that $P(n)$ will be true for $n = \frac{1}{2}$, because $\frac{1}{2}$ is not a natural number, so if you count $1, 2, 3, 4, \ldots$ you will never get to $\frac{1}{2}$. □

Imagine that you have some statement $P(n)$ about natural numbers that could be true or not for each natural number n. (For example, the statement $P(n)$ could be "$n(n+1)$ is even", or "n is even or odd", or "n is not both even and odd".) Suppose the following two facts are true:

I. The statement $P(n)$ is true for $n = 1$. (That is, $P(1)$ is true.)

II. Any time the statement $P(n)$ is true for one particular n, it follows that it is true for $n + 1$. (That is: if $P(n)$ is true then $P(n + 1)$ is true.)

The PMI says that, under these circumstances, $P(n)$ must be true for ***every*** natural number n.

THE PRINCIPLE OF MATHEMATICAL INDUCTION

Suppose $P(n)$ is any sentence in which n is an open variable.

Suppose, furthermore, that

I. $P(1)$ is true.

II. Any time $P(n)$ is true for one particular n, it follows that $P(n + 1)$ is true.)

Then $P(n)$ is true for every natural number n.

Let us say the same thing in formal language:

THE PRINCIPLE OF MATHEMATICAL INDUCTION (FORMAL LANGUAGE VERSION)

Suppose $P(n)$ is a sentence in which n is an open variable. Then

$$\Big(P(1) \wedge (\forall n \in \mathbb{N})(P(n) \Longrightarrow P(n+1))\Big) \Longrightarrow (\forall n \in \mathbb{N})P(n)\,. \tag{12.212}$$

12.3 The proof by induction that every natural number is even or odd and not both

We are going to use Theorems 22 (which says that 1 is good) and 27 (which says that goodness is passed on from each natural number n to its successor $n+1$).

We have already proved Theorem 22, but we have not proved Theorem 27, so we have to do it now.

Proof of Theorem 27.

Let n be an arbitrary natural number.

Assume that $\boxed{\boxed{n \text{ is good}}}$.

We are going to prove that $n+1$ is good.

Since n is good, n is even or odd, and n is not both even and odd.

Assume that $\boxed{n \text{ is even}}$.

Then n is not odd, because n is good.

It then follows from Theorem 23 that $n+1$ is odd.

It also follows from Theorem 23 that $n+1$ is not even. (Reason: If $n+1$ was even, then $(n+1)-1$ would be odd, that is, n would be odd. But n isn't odd[77].)

So $n+1$ is odd and $n+1$ is not even.

So $\boxed{n+1 \text{ is good}}$.

[77] Notice that this is a proof by contradiction

So $\boxed{n \text{ is even} \implies n+1 \text{ is good}}$.

Now assume that $\boxed{n \text{ is odd}}$.

Then n is not even, because n is good.

It then follows from Theorem 24 that $n+1$ is even.

It also follows from Theorem 24 that $n+1$ is not odd. (Reason: If $n+1$ was odd, then $(n+1)-1$ would be even that is, n would be even. But n isn't even[78].)

So $n+1$ is even $n+1$ is not odd.

So $\boxed{n+1 \text{ is good}}$.

So $\boxed{n \text{ is odd} \implies n+1 \text{ is good}}$.

Since we have "n is even $\vee$ n is odd", "n is even $\implies$ $n+1$ is good", and "n is odd $\implies$ $n+1$ is good", it follows from Rule $\vee_{prove}$ that $\boxed{\boxed{n+1 \text{ is good}}}$.

Since we have proved "$n+1$ is good" assuming "n is good", it follows from Rule $\implies_{prove}$ that

$$n \text{ is good} \implies n+1 \text{ is good}. \tag{12.213}$$

Since we have proved (12.213) for an arbitrary natural number n, it follows from Rule $\forall_{prove}$ that

$$(\forall n \in \mathbb{N})\Big(n \text{ is good} \implies n+1 \text{ is good}\Big). \tag{12.214}$$

We are now ready, finally, to prove the theorem that we had announced before, that every natural number is even or odd and not both.

We will prove this by induction.

[78] Another proof by contradiction!

THE FORMAT OF A PROOF BY INDUCTION

A proof by induction of a statement

$(\forall n \in \mathbb{N})XXXX$ should look like this:

Let $P(n)$ be the predicate XXXX.
Basis step. Proof of $P(1)$.
. . . . . .
$\boxed{P(1)}$.
Inductive step. We prove that
$(\forall n \in \mathbb{N})\Big(P(n) \Longrightarrow P(n+1)\Big)$.

Let $n \in \mathbb{N}$ be arbitrary. We want to prove $P(n) \Longrightarrow P(n+1)$.

Assume $P(n)$. We want to prove $P(n+1)$.
. . . . . .
$P(n+1)$.

So $P(n) \Longrightarrow P(n+1)$. [Rule $\Longrightarrow_{prove}$]

Hence $\boxed{(\forall n \in \mathbb{N})\Big(P(n) \Longrightarrow P(n+1)\Big)}$

[Rule $\forall_{prove}$]

We have completed the basis step and the inductive step. Hence it follows from the PMI that $(\forall n \in \mathbb{N})P(n)$.

That is, $(\forall n \in \mathbb{N})XXXX$. **Q.E.D.**

12.3.1 A remark on the importance of parentheses

PARENTHESES MATTER!!!

The sentence

$$(\forall n \in \mathbb{N})(P(n) \Longrightarrow P(n+1))\,. \qquad \text{(a)}$$

is not at all the same as the sentence

$$(\forall n \in \mathbb{N})P(n) \Longrightarrow P(n+1)\,. \qquad \text{(b)}$$

Sentence (a) says that the implication "$P(n) \Longrightarrow P(n+1)$" (that is, "$P$ is passed on from n to $n+1$") is true for every natural number n. So (a) says "every natural number passes on Property P to its successor".
Sentence (b) is totally different. It says: "if it is true that all natural numbers have P then $n+1$ has P". This is in fact meaningless, because n is an open variable.

12.3.2 Our first proof by induction: proof that every natural number is even or odd and not both

Theorem 28. *If n is a natural number, then*

1. *n is even (that is, $(\exists k \in \mathbb{Z})n = 2k$) or n is odd (that is, $(\exists k \in \mathbb{Z})n = 2k+1$);*

2. *n is not both even and odd.*

Proof. As we have been doing in previous sections, let us call an integer n "good" if n is even or odd and not both even and odd.

Let $P(n)$ be the sentence "n is good".

We want to prove that $(\forall n \in \mathbb{N})P(n)$.

Basis step. We have to prove $P(1)$, i.e., that 1 is good.

But we already know, from Theorem 22 that 1 good. So $\boxed{P(1) \text{ is true}}$, and this completes the basis step.

Inductive step. We have to prove that

$$(\forall n \in \mathbb{N})\big(P(n) \Longrightarrow P(n+1)\big)\,.$$

But we have alrady proved this, in Theorem 27, on page 232, which says precisely that goodness is passed on from an integer n to its successor $n+1$.

Since we have proved both that $P(1)$ and that

$$(\forall n \in \mathbb{N})\big(P(n) \Longrightarrow P(n+1)\big)\,,$$

it follows from the PMI that

$$(\forall n \in \mathbb{N})G(n)\,, \tag{12.215}$$

i.e., every natural number is good. **Q.E.D.**

Finally, we need to prove that every integer is good. It is very easy to prove that if $n \in \mathbb{Z}$ and n is good then $-n$ is good. (**YOU DO THIS.**)

Now, let n be an arbitrary integer. Then either $n \in \mathbb{N}$ or $-n \in \mathbb{N}$ or $n = 0$, by Basic Fact BFN4.

If $n \in \mathbb{N}$ then we already know that n is good.

If $-n \in \mathbb{N}$ then $-n$ is good, and then n is good as well.

So we have proved that the nonzero integers are good. If $n = 0$, then n is good as well because, for example, we alreadyknow that -1 is good, and goodness is passed on from each integer to its successor.

So we have proved that every integer is good. **Q.E.D.**

12.3.3 Proof that every integer is even or odd and not both

We now want to prove that every integer is good. That is, we want to prove:

Theorem 29. *If n is an integer, then*

1. *n is even (that is, $(\exists k \in \mathbb{Z})n = 2k$) or n is odd (that is, $(\exists k \in \mathbb{Z})n = 2k+1$);*
2. *n is not both even and odd.*

In order to prove this, we need two very simple theorems.

Theorem 30. *The integer 0 is even and not odd.*

Theorem 31. *If n is an integer then*

1. *If n is even then $-n$ is even.*
2. *If n is odd then $-n$ is odd.*
3. *If n is even and odd and not both, then $-n$ is even or odd and not both.*

Problem 57. ***Prove*** Theorems 30 and 31, using the theorems already proved in this section. □

Proof of Theorem 29.

As we have been doing in previous sections, let us call an integer n "good" if n is even or odd and not both even and odd.

We want to prove that every integer is good.

> Let $n \in \mathbb{Z}$ be arbitrary.
>
> Then either $n \in \mathbb{N}$, or $-n \in \mathbb{N}$, or $n = 0$.
>
> If $n \in \mathbb{N}$, then n is good by Theorem 28.
>
> If $-n \in \mathbb{N}$, then $-n$ is good by Theorem 28, and this implies that n is good by Theorem 31.
>
> If $n = 0$ then n is good by Theorem 30.
>
> So n is good. **Q.E.D.**

13 Examples of proofs by induction

13.1 Some divisibility theorems

Theorem 32. *If n is natural number, then $8^n - 5^n$ is divisible by 3.*

Proof. We want to prove that

$$(\forall n \in \mathbb{N}) 3|8^n - 5^n \,. \tag{13.216}$$

Let $P(n)$ be the predicate "$3|8^n - 5^n$

.

We want to prove that $(\forall n \in \mathbb{N})P(n)$.

We are going to prove this by induction.

Basis step:

We want to prove $P(1)$.

$P(1)$ aays "$3|8^1 - 5^1$".

And $8^1 = 8$, $5^1 = 5$, so $8^1 - 5^1 = 3$.

Therefore $3|8^1 - 5^1$, so $\boxed{P(1) \text{ is true}}$

Inductive step:

We want to prove $(\forall n \in \mathbb{N})(P(n) \Longrightarrow P(n+1))$.

Let $n \in \mathbb{N}$ be arbitrary.

Assume $P(n)$.
Then $3|8^n - 5^n$.
So we can write

$$8^n - 5^n = 3k \,, \;\; k \in \mathbb{Z} \,.. \tag{13.217}$$

Then

$$8 \times (8^n - 5^n) = 3 \times 8k \,. \tag{13.218}$$

So

$$8^{n+1} - 8 \times 5^n = 3 \times 8k \,, \tag{13.219}$$

and then

$$8^{n+1} = 8 \times 5^n + 3 \times 8k \,, \tag{13.220}$$

But $8 = 5 + 3$, so

$$8 \times 5^n = 5 \times 5^n + 3 \times 5^n = 5^{n+1} + 3 \times 5^n \,, \tag{13.221}$$

so

$$8^{n+1} = 5^{n+1} + 3 \times 5^n + 3 \times 8k \,, \tag{13.222}$$

and then

$$8^{n+1} = 5^{n+1} + 3(5^n + 8k) \,, \tag{13.223}$$

so that

$$8^{n+1} - 5^{n+1} = 3(5^n + 8k) \,, \tag{13.224}$$

Let $j = 5^n + 8k$. Then $j \in \mathbb{Z}$ and

$$8^{n+1} - 5^{n+1} = 3j \,. \tag{13.225}$$

Hence $3|8^{n+1} - 5^{n+1}$. That is, $\boxed{P(n+1)}$.

Therefore $\boxed{P(n) \Longrightarrow P(n+1)}$(by Rule $\Longrightarrow_{prove}$).

So $\boxed{(\forall n \in \mathbb{N})(P(n) \Longrightarrow P(n+1))}$ (by Rule $\forall_{prove}$).

This completes the inductive step.

Since we have proved $\boxed{P(1) \wedge (\forall n \in \mathbb{N})(P(n) \Longrightarrow P(n+1))}$, it follows from the PMI that $(\forall n \in \mathbb{N})P(n)$, that is,. $\boxed{(\forall n \in \mathbb{N})3|8^n - 5^n}$. **Q.E.D.**

Here are a few examples of theorems similar to Theorem 32

Theorem 33. *If n is natural number, then $11^n - 4^n$ is divisible by 7.*

Theorem 34. *If n is natural number, then $22^n - 10^n$ is divisible by 12.*

Theorem 35. *If n is natural number, then $31^n - 18^n$ is divisible by 13.*

Problem 58. ***Prove*** Theorem 33. □

Problem 59. ***Prove*** Theorem 34. □

Problem 60. ***Prove*** Theorem 35. □

Problem 61. *If, after reading the proof of Theorem 32 and solving Problems 58, 59, 60, you get the feeling that these are all the same thing,* ***try to prove*** *the following general theorem:*

Theorem 36. *If a, b are integers, then for every natural number n, $a^n - b^n$ is divisible by $a = b$.*

(This is done later, see Theorem 43 on page 260. But you shoud try to prove it by yourself before you look at the proof.) □

13.2 An inequality

Here is another example of a proof by induction.

Theorem 37. *If n is a natural number, then $2^n < n! + 3$.*

Proof. We want to prove that

$$(\forall n \in \mathbb{N}) 2^n < n! + 3\,. \tag{13.226}$$

Let $P(n)$ be the predicate "$2^n < n! + 3$".

We want to prove that $(\forall n \in \mathbb{N})P(n)$.

We are going to prove this by induction.

Basis step:

We want to prove $P(1)$.

$P(1)$ aays "$2^1 < 1! + 3$".

And $2^1 = 2$ and $1! + 3 = 4$.

Therefore $2^1 < 1! + 3$, so $\boxed{P(1) \text{ is true}}$

Inductive step:

We want to prove $(\forall n \in \mathbb{N})(P(n) \Longrightarrow P(n+1))$.

Let $n \in \mathbb{N}$ be arbitrary.

Assume $P(n)$. We want to prove $P(n+1)$.
Since $P(n)$ holds, we have

$$2^n < n! + 3\,. \tag{13.227}$$

Therefore, multiplying both sides of (13.227) by 2, we get

$$2^{n+1} < 2n! + 6\,. \tag{13.228}$$

On the other hand, $n + 1 = n - 1 + 2$, so

$$(n+1)! = (n+1)n! = (n-1)n! + 2n!\,. \tag{13.229}$$

We are going to treat separately the cases $n \geq 3$ and $n < 3$.
Assume that $n \geq 3$.

Then $n-1 \geq 2$ and $n! \geq 6$, so $(n-1)n! \geq 12$ and *a fortiori* $(n-1)n! > 3$.

$*$ Since $(n+1)! = (n-1)n! + 2n!$, and $(n-1)n! > 3$, we have $(n+1)! > 2n! + 3$, that is

$$2n! + 3 < (n+1)! \,. \tag{13.230}$$

Since $2^{n+1} < 2n! + 6$, we have

$$\begin{aligned} 2^{n+1} &< 2n! + 6 \\ &= 2n! + 3 + 3 \\ &< (n+1)! + 3 \,, \end{aligned}$$

so $2^{n+1} < (n+1)! + 3$.

That is, $\boxed{P(n+1) \text{ holds.}}$

We now consider the case when $n < 3$.

Assume that $n < 3$.

Then $n = 1$ or $n = 2$,

If $n = 1$ then $P(n+1)$ says $2^2 < 2! + 3$, that is $4 < 5$. So $P(n+1)$ is true.

If $n = 2$ then $P(n+1)$ says $2^3 < 3! + 3$, that is $8 < 9$. So $P(n+1)$ is true.

So in both cases $\boxed{P(n+1) \text{ holds.}}$

We have proved that $P(n+1)$ holds in both case, when $n \geq 3$ and when $n < 3$. So $\boxed{\boxed{P(n+1).}}$

Therefore $\boxed{P(n) \Longrightarrow P(n+1)}$(by Rule $\Longrightarrow_{prove}$).

So $\boxed{(\forall n \in \mathbb{N})(P(n) \Longrightarrow P(n+1))}$ (by Rule $\forall_{prove}$).

This completes the inductive step.

Since we have proved $\boxed{P(1) \wedge (\forall n \in \mathbb{N})(P(n) \Longrightarrow P(n+1))}$, it follows from the PMI that $(\forall n \in \mathbb{N})P(n)$, that is,. $\boxed{(\forall n \in \mathbb{N})2^n < n! + 3}$. **Q.E.D.**

Problem 62.

1. ***Prove*** that if n is a natural number then $3^n < n! + 124$.
2. Is it true that if n is a natural number then $3^n < n! + 123$?

13.3 More inequalities, with applications to the computation of some limits

Let us use induction to prove an inequality:

Theorem 38. *If x is a positive real number, and n is a natural number, then*

$$(1+x)^n \geq 1+nx\,. \tag{13.231}$$

Proof. We want to prove that

$$(\forall x \in \mathbb{R})(\forall n \in \mathbb{N})\Big(x > 0 \Longrightarrow (1+x)^n \geq 1+nx\Big)\,. \tag{13.232}$$

Let x be an arbitrary real number.

We want to prove that

$$(\forall n \in \mathbb{N})\Big(x > 0 \Longrightarrow (1+x)^n \geq 1+nx\Big)\,. \tag{13.233}$$

We prove this by induction.

Let $P(n)$ be the predicate "$x > 0 \Longrightarrow (1+x)^n \geq 1+nx$".

Base step. We have to prove $P(1)$.

But $P(1)$ says "$x > 0 \Longrightarrow 1+x \geq 1+x$", and this implication is obviously true, because its conclusion is true.

So $P(1)$ is true, and we are done with the base case.

Inductive step. We have to prove

$$(\forall n \in \mathbb{N})(P(n) \Longrightarrow P(n+1))\,. \tag{13.234}$$

Let n be an arbitrary natural number. We want to prove that $P(n) \Longrightarrow P(n+1)$.

Assume $P(n)$.

Then

$$x > 0 \Longrightarrow (1+x)^n \geq 1+nx\,. \tag{13.235}$$

We want to prove

$$x > 0 \Longrightarrow (1+x)^{n+1} \geq 1+(n+1)x\,. \tag{13.236}$$

Assume $x > 0$.

Then it follows from (13.235) (by Rule $\Longrightarrow_{use}$) that

$$(1+x)^n \geq 1 + nx\,. \tag{13.237}$$

Multiplying both sides of (13.237) by $1+x$ (which is possible because $1 + x > 0$), we get

$$(1+x)^{n+1} \geq (1+x)(1+nx)\,. \tag{13.238}$$

But

$$\begin{aligned}(1+x)(1+nx) &= 1 + x + nx + nx^2 \\ &= 1 + (n+1)x + nx^2 \\ &\geq 1 + (n+1)x\,.\end{aligned}$$

(The fact that $1 + (n+1)x + nx^2 \geq 1 + (n+1)x$ follows because $nx^2 \geq 0$ and then, adding $1+(n+1)x$ to both sides, we get $1 + (n+1)x + nx^2 \geq 1 + (n+1)x$.)

So

$$(1+x)^{n+1} \geq 1 + (n+1)x\,. \tag{13.239}$$

Since we proved (13.239) under the assumption that $x > 0$, it follows that

$$x > 0 \Longrightarrow (1+x)^{n+1} \geq 1 + (n+1)x\,. \tag{13.240}$$

That is, $P(n+1)$ holds.

Since we have proved $P(n+1)$ assuming $P(n)$, Rule $\Longrightarrow_{prove}$ allows us to conclude that $P(n) \Longrightarrow P(n+1)$.

So we have proved $P(n) \Longrightarrow P(n+1)$ for arbitrary $n \in \mathbb{N}$, Rule $\forall_{prove}$ allows us to conclude that (13.234) holds.

This completes the inductive step.

Since we have also proved $P(1)$, we can use the PMI to conclude that (13.233) holds, i.e., that

$$(\forall n \in \mathbb{N})\Big(x > 0 \Longrightarrow (1+x)^n \geq 1 + nx\Big)\,. \tag{13.241}$$

Since we have proved for an arbitrary real number x, we can conclude that

$$(\forall x \in \mathbb{R})(\forall n \in \mathbb{N})\Big(x > 0 \Longrightarrow (1+x)^n \geq 1 + nx\Big)\,, \tag{13.242}$$

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

Problem 63. In the proof of Theorem 38, we translated the statement to be proved into formal language as Formula (13.232) and then followed the rules of logic, plus the PMI, to prove it.

Suppose instead that we had translated the statement of Theorem 38 in a different way, as

$$(\forall n \in \mathbb{N})(\forall x \in \mathbb{R})\Big(x > 0 \Longrightarrow (1+x)^n \geq 1 + nx\Big)\,. \tag{13.243}$$

1. ***Prove that this translation is equivalent to Formula (13.232)***, as a matter of pure logic. That is, prove that no matter what the 2-variable predicate $A(x,n)$ is, and what the sets S, T are, the formulas

   $$(\forall x \in S)(\forall n \in T)A(x,n)$$

   and

   $$(\forall n \in T)(\forall x \in S)A(x,n)$$

   are equivalent. (Two formulas U, V are equivalent if $U \Longleftrightarrow V$ is true.)

2. ***Write a different proof*** of Theorem 38, using the translation (13.243) instead of (13.232).

Problem 64. By looking carefully at the proof of Theorem 38, ***prove*** the following stronger result:

Theorem 39. *If $x \in \mathbb{R}$ and $x \geq -1$, and n is a natural number, then*

$$(1+x)^n \geq 1 + nx\,. \tag{13.244}$$

With a little bit more work, it is possible to prove a result stronger than Theorem 38:

Theorem 40. *If x is a nonnegative real number, and n is a natural number, then*

$$(1+x)^n \geq 1 + nx + \frac{n(n-1)}{2}x^2\,. \tag{13.245}$$

Proof.

YOU DO THIS ONE.

HINT. Just repeat the proof of Theorem 38 up to the point when you multiply by $1+x$, and at that point keep the x^2 term. □

Problem 65. ***Prove*** Theorem 40. □

13.3.1 An application of Theorem 40: computing $\lim_{n\to\infty} \sqrt[n]{n}$

In this section we use the notion of "limit of a sequence". All you need to know about limits of seuences is the following sandwiching theorem": If $\{a_n\}_{n=1}^{\infty}$, $\{b_n\}_{n=1}^{\infty}$, and $\{c_n\}_{n=1}^{\infty}$, are sequences of real numbers such that $a_n \le b_n \le c_n$ for every $n \in \mathbb{N}$, and L is a real number such that

$$\lim_{n\to\infty} a_n = L \quad \text{and} \quad \lim_{n\to\infty} c_n = L\,,$$

then $\lim_{n\to\infty} b_n$.

Let us prove that

$$\lim_{n\to\infty} \sqrt[n]{n} = 1\,. \tag{13.246}$$

Define

$$\alpha_n = \sqrt[n]{n} - 1\,.$$

To prove (13.246), we have to prove that

$$\lim_{n\to\infty} \alpha_n = 0\,. \tag{13.247}$$

It is clear that $\alpha_n \ge 0$. (Reason: $\sqrt[n]{n} \ge 1$, because if $\sqrt[n]{n}$ was < 1, it would follow that $\left(\sqrt[n]{n}\right)^n < 1$, but $\left(\sqrt[n]{n}\right)^n = n$, and $n \ge 1$.)

Also, $1 + \alpha_n = \sqrt[n]{n}$, so

$$(1+\alpha_n)^n = n\,. \tag{13.248}$$

Using the inequality of Theorem 40, we get

$$(1+\alpha_n)^n \ge 1 + n\alpha_n + \frac{n(n-1)}{2}\alpha_n^2\,. \tag{13.249}$$

So

$$\begin{aligned} n &= (1+\alpha_n)^n \\ &\ge 1 + n\alpha_n + \frac{n(n-1)}{2}\alpha_n^2 \\ &\ge \frac{n(n-1)}{2}\alpha_n^2\,. \end{aligned}$$

Hence

$$n \geq \frac{n(n-1)}{2}\alpha_n^2\,,$$

so

$$1 \geq \frac{n-1}{2}\alpha_n^2\,,$$

and then

$$\alpha_n^2 \leq \frac{2}{n-1}\,,$$

so

$$\alpha_n \leq \sqrt{\frac{2}{n-1}}\,.$$

Hence the numbers α_n satisfy

$$0 \leq \alpha_n \leq \sqrt{\frac{2}{n-1}}\,.$$

So the α_n are 'sandwiched" between two sequences that converge to 0. Hence $\lim_{n\to\infty} \alpha_n = 0$ by the sandwiching theorem.

Hence (13.246 is proved.

13.4 Some formulas for sums

In this section we use the notation "$\sum_{k=1}^{n} a_k$" for "$a_1 + a_2 + \cdots + a_n$". (A precise definition of "$\sum_{k=1}^{n} a_k$", without using $\cdots$, is given in section 13.5.3 on page 257.)

Theorem 41. *If n is an arbitrary natural number, then*

$$\sum_{k=1}^{n} k = \frac{n(n+1)}{2}\,. \tag{13.250}$$

(That is, $1 + 2 + \cdots + n = \frac{n(n+1)}{2}$.)

Proof. Let $P(n)$ be the statement "$\sum_{k=1}^{n} = \frac{n(n+1)}{2}$".

We prove $(\forall n \in \mathbb{N})P(n)$ by induction.

Base step. $P(1)$ says "$1 = \frac{1(1+1)}{2}$", which is obviously true. So $P(1)$ is true.

Inductive step.

We prove $(\forall n \in \mathbb{N})(P(n) \Longrightarrow P(n+1))$.

Let n be an arbitrary natural number.

Assume that $P(n)$ is true.
Then $\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$.
Therefore

$$\begin{aligned}
\sum_{k=1}^{n+1} k &= (\sum_{k=1}^{n} k) + (n+1) \\
&= \frac{n(n+1)}{2} + (n+1) \\
&= (n+1)\left[\frac{n}{2} + 1\right] \\
&= (n+1) \times \frac{n+2}{2} \\
&= \frac{(n+1)(n+2)}{2}\,.
\end{aligned}$$

So

$$\sum_{k=1}^{n+1} k = \frac{(n+1)(n+2)}{2}\,.$$

That is, $P(n+1)$ holds.

We have proved $P(n+1)$ assuming $P(n)$. Hence $\boxed{P(n) \Longrightarrow P(n+1)}$.

We have preved $P(n) \Longrightarrow P(n+1)$ for an arbitrary natural number n. Therefore $(\forall n \in \mathbb{N})(P(n) \Longrightarrow P(n+1))$, which completes the inductive step.
Hence, by the PMI, $(\forall n \in \mathbb{N})P(n)$, that is,

$$(\forall n \in \mathbb{N}) \sum_{k=1}^{n} k = \frac{n(n+1)}{2}\,.$$

Q.E.D.

Using the same method, many other formulas for sums can be proved. Here is an exmaple of a rather remarkable one:

Theorem 42. *If n is a natural number, then*

$$\sum_{k=1}^{n} k^3 = \left[\frac{n(n+1)}{2}\right]^2, \tag{13.251}$$

that is:

$$1^3 + 2^3 + 3^3 + 4^3 + \cdots + n^3 = \left[\frac{n(n+1)}{2}\right]^2.$$

Proof. **YOU DO THIS ONE.**

Problem 66.

1. ***Compute*** the sum $\sum_{k=1}^{n} k^3$ for $n = 1, 2, 3, 4, 5$ and 6.
2. ***Verify*** that in each case the sum you got is a perfect square (i.e., the square of an integer).
3. ***Prove*** Theorem 42. □

Problem 67.

1. ***Compute*** the sum $\sum_{k=1}^{n} k^2$ for $n = 1, 2, 3, 4, 5$ and 6.
2. ***Verify*** that in each case the sum you got agrees with the formula

$$\sum_{k=1}^{n} k^2 = \frac{n + 3n^2 + 2n^3}{6}. \tag{13.252}$$

3. ***Prove*** that Formula (13.252) holds for every natural number n. □

Problem 68.

1. ***Compute*** the sum $\sum_{k=1}^{n} k$ for $n = 1, 2, 3, 4, 5$ and 6.
2. ***Verify*** that in each case the sum you got agrees with the formula

$$\sum_{k=1}^{n} k = \frac{n(n+1)}{2}. \tag{13.253}$$

3. ***Prove*** that Formula (13.253) holds for every natural number n. □

Problem 69.

1. ***Compute*** the sum $\sum_{k=1}^{n}(2k-1)$ for $n = 1, 2, 3, 4, 5$ and 6.
2. ***Verify*** that in each case the sum you got agrees with the formula

$$\sum_{k=1}^{n}(2k-1) = n^2\,. \tag{13.254}$$

3. ***Prove*** that Formula (13.254) holds for every natural number n. □

Problem 70. ***Figure out*** a formula for the sum

$$\sum_{k=1}^{n}(2k-1)^2\,, \tag{13.255}$$

and ***prove*** that your formula holds for every natural number n. □

Problem 71. ***Figure out*** a formula for the sum

$$\sum_{k=1}^{n}(4k+3)^3\,, \tag{13.256}$$

and ***prove*** that your formula holds for every natural number n. □

13.5 Inductive definitions

In an earlier set of lectures, we defined "x^2", for a real number x, to mean "$x.x$". And we can define "x^3" to mean "$(x.x).x$", or, if you prefer, "$x^2.x$". But how can we define "x^n" for an arbitrary natural number n? One possibility would be to write something like this

$$x^n = \underbrace{x \times x \times \cdots \times x}_{n \text{ times}}$$

Similarly, we would like to define the "factorial" $n!$ of a natural number n by the formula

$$n! = 1 \times 2 \times 3 \times \cdots \times n\,.$$

And we would like to define summations such as

$$1 + 2 + 3 + \cdots + n$$

or

$$1^2 + 2^2 + 3^2 + \cdots + n^2\,,$$

or products such that

$$2 \times 4 \times 6 \times 8 \times \cdots \times 200\,.$$

With this notation, if we want to talk about the product of the first 20 prime numbers, i.e., the number

$$2\times3\times5\times7\times11\times13\times17\times19\times23\times29\times31\times37\times41\times43\times47\times53\times59\times61\times67\times71\,,$$

we could write

$$2 \times 3 \times \cdots \times 71\,. \tag{13.257}$$

But this is very unclear. I do not know what "$\cdots$" means, precisely (and if you think you do, please tell me!). For example, in the expression (13.257), how on Earth are we supposed to know which numbers should go in place of the $\cdots$? Take a simple example of a similar situation: suppose I write

$$3 \times 5 \times 7 \times \cdots \times 71\,. \tag{13.258}$$

Is this supposed to be "the product of all odd numbers from 3 to 71", or "the product of all prime numbers from 3 to 71", or "the product of all the odd numbers from 3 to 71 that do not end in a 9", or what?

Next, let us look at another example: suppose I write

$$1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, 377, \ldots.$$

What is the next number, after 377? Well, if you have guessed the pattern, then you will probably guess that each number, after the first two, is the sum of the two preceding ones, so what comes after 377 is $233 + 377$, that is, 610. But, why couldn't the pattern be this:

- Start with 1, and then another 1.
- Then each number is obtained by adding the two preceding ones.
- Yo go on like this until you get to 377, and then you switch to a different rule: each number is obtained by adding 100 to the previous one.

This is a perfectly legitimate rule for generating a sequence of numbers, and if you use this rule then the numbers that come after 377 are 477, 577, and so on. If you say "that's not a true pattern", then I will ask you to tell me what you mean by "a true pattern", and I will also ask "Why not? What do you mean by 'pattern'?". "Why is this not a true pattern?".

One last example. If I write

$$27, 82, 41, 124, 61, 184, 92, 46, \cdots$$

what comes next? I'll let you think about this one.

The fact is: in general, "$\cdots$" is meaningless. So in mathematics we just do not use it.

And, in any case, once we develop fully our way of writing all of mathematics formally (that is, with formulas and no words), the symbol "$\cdots$" will not be there in the list of symbols we can use. So we do not want to use "$\cdots$" at all.

What we are going to do instead is use ***inductive definitions***.

13.5.1 The inductive definition of powers of a real number

The way to define the power "x^n" correctly is by means of an inductive definition: we first define x^1 to be x, and then define x^{n+1} to be $x^n.x$, for every n. That is, we write:

Definition 18. ***(Inductive definition of positive integer powers of a real number)*** *For all $a \in \mathbb{R}$, we set*

$$\begin{aligned} a^1 &= a\,, \\ a^{n+1} &= a^n.a \quad \text{for } n \in \mathbb{N}\,. \end{aligned}$$

We also set $a^0 = 1$. □

Using this definition, we can write down what a^n is for any n.

Suppose, for example, that we want to know what a^5 is. By the second line of our inductive definition of a^n,

$$a^5 = a^4.a.$$

This answers our question about a^5, in terms of a^4. And what is a^4? Again, using the second line of the inductive definition, we find

$$a^4 = a^3.a.$$

So

$$a^5 = ((a^3).a).a.$$

And what is a^3? Once again, we can use the second line of the inductive definition, and find

$$a^3 = a^2.a$$

So

$$a^5 = (((a^2).a).a).a.$$

One more step yields

$$a^2 = a^1.a\,,$$

so

$$a^5 = (((a^1.a).a).a).a.$$

And, finally, the first line of the inductive definition, tells us that $a^1 = a$, so we end up with

$$a^5 = (((a.a).a).a).a.$$

Furthermore, since multiplication of real numbers has the associative property, we can omit the parentheses and just write:

$$a^5 = a.a.a.a.a.$$

13.5.2 The inductive definition of the factorial

The "factorial" of a natural number n is supposed to be the product $1 \times 2 \times 3 \times \cdots \times n$. That is, the factorial of n is the product of all the natural numbers from 1 to n. Here is the inductive definition:

Definition 19. *The factorial of a natural number n is the number $n!$ given by*

$$1! = 1\,, \tag{13.259}$$

$$(n+1)! = n! \times (n+1) \quad \text{for } n \in \mathbb{N}\,. \tag{13.260}$$

In addition, we define

$$0! = 1\,,$$

so $n!$ is defined for every nonnegative integer n. □

Example 61. Let us compute 7! using the inductive definition. Using (13.260) we get $7! = 7 \times 6!$. Then using (13.260) again we get $6! = 6 \times 5!$, so $7! = 7 \times 6 \times 5!$. Continuing in the same way we get $5! = 5 \times 4!$, so $7! = 7 \times 6 \times 5 \times 4!$, and then $4! = 4 \times 3!$, so $7! = 7 \times 6 \times 5 \times 4 \times 3!$. Then $3! = 3 \times 2!$, so $7! = 7 \times 6 \times 5 \times 4 \times 3 \times 2!$. And $2! = 2 \times 1!$, so $7! = 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1!$. Finally, (13.259) tells us that $1! = 1$, so we end up with

$$7! = 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1\,,$$

which is of course what 7! is supposed to be. □

13.5.3 The inductive definition of summation.

Definition 20. *Suppose we have a natural number n, and a list*

$$\mathbf{a} = (a_1, a_2, \ldots, a_n)$$

of n real numbers. We define the sum (or summation) of the list $\mathbf{a}$ (also called the sum of the a_j for j from 1 to n) to be the number $\sum_{j=1}^{n} a_j$ determined as follows:

$$\begin{aligned} \sum_{j=1}^{1} a_j &= a_1\,, \\ \sum_{j=1}^{n+1} a_j &= \Big(\sum_{j=1}^{n} a_j\Big) + a_{n+1} \quad \text{for} \quad n \in \mathbb{N}\,. \end{aligned}$$

And we also define $\sum_{j=1}^{0} a_j = 0$.

Example 62. Let us compute $\sum_{j=1}^{5} j^2$. We have

$$\begin{aligned} \sum_{j=1}^{5} j^2 &= \Big(\sum_{j=1}^{4} j^2\Big) + 5^2 \\ &= \left(\Big(\sum_{j=1}^{3} j^2\Big) + 4^2\right) + 5^2 \\ &= \Big(\sum_{j=1}^{3} j^2\Big) + 4^2 + 5^2 \\ &= \Big(\sum_{j=1}^{2} j^2\Big) + 3^+4^2 + 5^2 \\ &= \Big(\sum_{j=1}^{1} j^2\Big) + 2^2 + 3^+4^2 + 5^2 \\ &= 1^2 + 2^2 + 3^+4^2 + 5^2 \\ &= 1 + 4 + 9 + 16 + 25 \\ &= 55\,. \end{aligned}$$

13.5.4 Inductive definition of product.

Definition 21. *For a natural number n, and a list $\mathbf{a} = (a_1, a_2, \ldots, a_n)$ of n real numbers, we define the product of the a_j for j from 1 to n to be the*

number $\prod_{j=1}^{n} a_j$ *determined as follows:*

$$\begin{aligned} \prod_{j=1}^{1} a_j &= a_1\,, \\ \prod_{j=1}^{n+1} a_j &= \Big(\prod_{j=1}^{n} a_j\Big) \times a_{n+1} \quad \text{for} \quad n \in \mathbb{N}\,. \end{aligned}$$

And we also define $\prod_{j=1}^{0} a_j = 1$.

Example 63. If you compare the inductive definition of a product with the inductive definition of the factorial, you can easily see that

$$n! = \prod_{j=1}^{n} j \qquad \text{for every } n \in \mathbb{N}\,.$$

13.5.5 A simple example of a proof by induction using inductive definitions

Here is a simple example of a proof of an inequality by induction. Notice how the proof uses the notion of "n-th power" of a real number exactly in the form of the inductive definition.

Proposition 1. *For all* $n \in \mathbb{N}$, $n < 2^n$.

Proof.
Let $P(n)$ be the statement "$n < 2^n$".
We are going to prove

$$(\forall n \in \mathbb{N})P(n) \tag{13.261}$$

by induction

Basis step. $P(1)$ is the statement "$1 < 2^1$". But $2^1 = 2$ by the inductive definition, so $P(1)$ says "$1 < 2$" which is clearly true. So $\boxed{P(1)}$ is true.

Inductive step. We want to prove that

$$(\forall n \in \mathbb{N})(P(n) \Longrightarrow P(n+1))\,. \tag{13.262}$$

Let n be an arbitrary natural number.

We want to prove that $P(n) \Longrightarrow P(n+1)$.

Assume $P(n)$.

Then $n < 2^n$.

So $2n < 2^n \times 2 = 2^{n+1}$.

But $1 \leq n$, because n is a natural number. (Precisely: if $n = 1$ then $1 = n$, so $1 \leq n$. And if $n \neq 1$ then ny Basic Fact BFZ9, $n - 1 \in \mathbb{N}$, so $1 < n$, and then $1 \leq n$.)

So $n + 1 \leq n + n$, i.e., $n + 1 \leq 2n$.

Therefore $n + 1 < 2^{n+1}$.

So $P(n + 1)$ is true.

Since we have proved $P(n + 1)$ assuming $P(n)$, we can conclude that $P(n) \Longrightarrow P(n + 1)$.

Since we have proved $P(n) \Longrightarrow P(n + 1)$ for arbitrary n, it follows that (13.262) holds.
So we have completed the basis step and the inductive step, and then the PMI tells us that (13.261 holds, that is, that $(\forall n \in \mathbb{N}) n < 2^n$. **Q.E.D.**

13.5.6 Another simple example of a proof by induction using inductive definitions

Here is a slightly more involved example of a proof of an inequality by induction. Notice how the proof uses the notion of "n-th power" of a real number and the notion of "factorial" exactly in the form of their inductive definitions.

We would like to prove the inequality "$2^n < n!$". This, however, isn't true for every natural number n. (For example, it is not true if $n = 1$ or $n = 2$ or $n = 3$.) But it is true for $n \geq 4$.

Proposition 2. *For all $n \in \mathbb{N}$, if $n \geq 4$ then $2^n < n!$.*

Proof.
Let $P(n)$ be the statement "$2^n < n!$".
We are going to prove

$$(\forall n \in \mathbb{N})(n \geq 4 \Longrightarrow P(n)) \,. \tag{13.263}$$

by induction. And we will start the induction at 4 rather than 1.

Basis step. $P(4)$ is the statement "$2^4 < 4!$". But $2^4 = 16$, and $4! = 24$. So $P(1)$ says "$16 < 24$', which is clearly true. So $\boxed{P(4)}$ is true.

Inductive step. We want to prove that

$$(\forall n \in \mathbb{N})\Big(n \geq 4 \Longrightarrow (P(n) \Longrightarrow P(n+1))\Big)\,. \tag{13.264}$$

Let n be an arbitrary natural number such that $n \geq 4$..
We want to prove that $P(n) \Longrightarrow P(n+1)$.

Assume $P(n)$.
Then $2^n < n!$.
So $2 \times 2^n < 2n!$.
But $2 \times 2^n = 2^{n+1}$.
Hence $2^{n+1} < 2n!$.
Also, $2 < n+1$.
So $2n! < (n+1)n!$.
But $(n+1)n! = (n+1)!$ by the inductive definitionof "factorial".

Therefore $2n! < (n+1)!$.
So, finally, $2^{n+1} < (n+1)!$.
So $P(n+1)$ is true.

Since we have proved $P(n+1)$ assuming $P(n)$, we can conclude that $P(n) \Longrightarrow P(n+1)$.

Since we have proved $P(n) \Longrightarrow P(n+1)$ for arbitrary n, it follows that (13.264) holds.
So we have completed the basis step and the inductive step, and then the PMI tells us that (13.263) holds, that is, that $(\forall n \in \mathbb{N})\Big(n \geq 4 \Longrightarrow (2^n < n!)\Big)$. **Q.E.D.**

13.5.7 Another simple example: divisibility by 3, 9, and 11

Let us prove

Theorem 43. *If a, b are arbitrary integers, then for every nonnegative integer*[79] *n the integer $a^n - b^n$ is divisible by $a - b$.*

Example 64. Here are some examples of what the theorem says:

1. Take $a = 8$, $b = 3$. Then the theorem says that $8^n - 3^n$ is divisible by 5 for every n. (And you can check this. For example, $8^3 = 512$, and $3^3 = 27$, so $8^3 - 3^3 = 512 - 27 = 495$, which is indeed divisible by 5.

[79] Recall that the ***nonnegative integers*** are the natural numbers as well as zero.

2. Take $a = 10$, $b = 1$. Then the theorem says that $10^n - 1$ is divisible by 9, and you can check this. (For example, $10^1 - 1 = 9$, $10^2 - 1 = 99$, $10^3 - 1 = 999$, $10^4 - 1 = 9,999$, and so on.)

3. Take $a = 10$, $b = -1$. Then the theorem says that $10^n - (-1)^n$ is divisible by 11. And you can check this: $10-(-1) = 11$, $10^2-(-1)^2 = 99$, $10^3 - (-1)^3 = 1,001$, $10^4 - (-1)^4 9,999$, and all these are divisible by 11. □

Proof.

Let a, b be arbitrary integers.

We will prove that

$$(\forall n \in \mathbb{N}) a - b | a^n - b^n \,, \tag{13.265}$$

and also that "$a - b | a^n - b^n$" is true for $n = 0$.

First we prove (13.265) by induction.

Let $P(n)$ be the statement[80] "$a - b$ divides $a^n - b^n$".

Basis Step. $P(1)$ says "$a - b$ divides $a - b$", which is obviously true.

This completes the basis step.

Inductive Step. We want to rpove

Inductive step. We want to prove that

$$(\forall n \in \mathbb{N})(P(n) \Longrightarrow P(n+1)) \,. \tag{13.266}$$

Let n be an arbitrary natural number.

We want to prove that $P(n) \Longrightarrow P(n+1)$.

Assume $P(n)$.

Then $a - b$ divides $a^n - b^n$.

So we may pick an integer k such that

$$a^n - b^n = (a - b)k \,. \tag{13.267}$$

[80] We do not have to worry about the quesion "who are a and b?", because we have fixed a and b earlier. They are fixed integers. Arbitrary, but fixed.

Then

$$\begin{aligned} a^{n+1} - b^{n+1} &= a^{n+1} - ab^n + ab^n - b^{n+1} \\ &= aa^n - ab^n + ab^n - bb^n \\ &= a(a^n - b^n) + (a-b)b^n \\ &= a(a-b)k + (a-b)b^n \\ &= (a-b)\big(ak + b^n\big)\,. \end{aligned}$$

Hence $a^{n+1} - b^{n+1} = (a-b)\big(ak + b^n\big)$.

Clearly, $ak + b^n$ is an integer[81].

Therefore $a - b$ divides $a^{n+1} - b^{n+1}$.

So $P(n+1)$ is true.

Since we have proved $P(n+1)$ assuming $P(n)$, we can conclude that $P(n) \Longrightarrow P(n+1)$.

Since we have proved $P(n) \Longrightarrow P(n+1)$ for arbitrary n, it follows that (13.266) holds.

So we have completed the basis step and the inductive step, and then the PMI tells us that (13.265 holds, that is, that if n is an arbitrary natural number, then $a - b$ divides $a^n - b^n$.

This almost completes our proof. But there is a minor missing detail: we also have to prove that $a - b$ divides $a^n - b^n$ when $n = 0$.

But if $n = 0$ then $a^n - b^n$ is equal to zero, because the inductive definition of the powers tells us that $a^0 = 1$ and $b^0 = 1$.

And 0 is divisible by any integer.

So $a - b$ divides $a^n - b^n$ also when $n = 0$.

We have now proved that $a - b | a^n - b^n$ for every nonnegative integer n.

And this has been proved for arbitrary integers a, b. So our proof is complete. **Q**.**E**.**D**.

[81] Strictly speaking even a stupid, trivial, obvious statement like this needs proof. On the other hand, it is so obvous that nobody would actually insult the reader's intelligence by putting in the proof. On the other hand, at this point we are just getting started with proofs, so you shousl knwo how to prove this. So I am going to ask you to write down the proof, as a homework problem. ***Sorry!***.

Problem 72.

1. ***Provide a detailed proof*** of the step that we skipped in the proof of Theorem 43, namely, that $ak + b^n$ is an integer. (This will require proving that if $b \in \mathbb{Z}$ then $b^n \in \mathbb{Z}$ for every nonnegative integer n, and the only way to do that is by induction, using the inductive definition of the powers.)

2. ***Provide an alternative proof*** of Theorem 43, in which you do not treat separately the cases $n \in \mathbb{N}$ and $n = 0$, but do the whole thing in one swoop, using the PMI starting at 0 rather than at 1.

3. ***Explain*** how you would answer the following objection that somebody studying these notes might raise: *In the theorem, you do not assume that $a \neq b$, and you talk about "divisibility by $a - b$". But if $a = b$ then $a - b$ is zero, and we cannot divide by zero, so how come you allow a to be equal to b? How can you say that "0 is divisible by 0", given that $\frac{0}{0}$ is not defined?* □

Problem 73. One of the consequences of Theorem 43 is that $10^n - 1$ is divisible by 9 for each nonnegative integer n. So, for example, if you look at the number 438, and let $s = 4 + 3 + 8$, so $s = 15$, it follows that $438 - s$ is divisible by 9, because:

$$\begin{aligned} & 438 - s \\ = \;& 4 \times 100 + 3 \times 10 + 4 \times 1 - (4 + 3 + 8) \\ = \;& 4 \times 10^2 - 4 + 3 \times 10 - 3 + 4 \times 1 - 1 \\ = \;& 4 \times (10^2 - 1) + 3 \times (10 - 1) + 4 \times (1 - 1)\,, \end{aligned}$$

which is clearly divisible by 9.

1. ***Explain*** how this fact leads to the following two divisibility criteria:

   Criterion for divisibilitly by 9***:*** A natural number n is divisible by 9 if and only if the sum of its decimal figures is divisible by 9. (For example: $572,265$ is divisible by 9 because $5 + 7 + 2 + 2 + 6 + 5 = 27$, which is divisible by 9. And $772,265$ is not divisible by 9 because $7 + 7 + 2 + 2 + 6 + 5 = 29$, which is not divisible by 9.)

   Criterion for divisibilitly by 3***:*** A natural number n is divisible by 3 if and only if the sum of its decimal figures is divisible by 3. (For example: $572,265$ is divisible by 3 because

$5 + 7 + 2 + 2 + 6 + 5 = 27$, which is divisible by 3. And $772,265$ is not divisible by 3 because $7 + 7 + 2 + 2 + 6 + 5 = 29$, which is not divisible by 3.)

2. Explain, in a similar way, how the fact that $10^n - (-1)^n$ is divisible by 11 leads to the following divisibility criterion:

   Criterion for divisibilitly by 11***:*** A natural number n is divisible by 11 if and only if the alternating sum[82] of its decimal figures is divisible by 11. (For example: $572,473$ is divisible by 11 because $5-7+2-4+7-3 = 0$, which is divisible by 11. And $772,463$ is not divisible by 11 because $7 - 7 + 2 - 4 + 6 - 3 = 1$, which is not divisible by 11.) □

13.5.8 Some problems

Problem 74. ***Prove***, using the inductive definition of the powers a^n, that

1. $(\forall a \in \mathbb{R})(\forall b \in \mathbb{R})(\forall n \in \mathbb{N})(ab)^n = a^n b^n$,
2. $(\forall a \in \mathbb{R})(\forall m \in \mathbb{N})(\forall n \in \mathbb{N})a^{m+n} = a^m a^n$. □

Problem 75. ***Prove***, using the inductive definition of summation, that if $n \in \mathbb{N}$ and $(a_1, a_2, \ldots, a_n)$ and
$(b_1, b_2, \ldots, b_n)$, are finite lists of natural numbers of length n, then

$$\sum_{k=1}^{n} a_k + \sum_{k=1}^{n} b_k = \sum_{k=1}^{n} (a_k + b_k)\,. \tag{13.268}$$

[82] That is, the sum with alternating signs: first figure minus second figureplus third figure minus fourth figure, etc, etc.

14 Other forms of induction

14.1 Induction with a different starting point (sometimes called "generalized induction")

The PMI says that, if a property is true of 1, and is passed on to the right, so each natural number n passes it on to its successor $n+1$, then the property will hold of all the numbers that we reach by counting starting at 1.

It is clear that the same thing should be true if we start counting at some other starting point s_*, that is, some other integer such as, for example, 3, or 7, or 0, or -5, or -372. The general result is the following rather trivial theorem:

THE PRINCIPLE OF MATHEMATICAL INDUCTION WITH A GENERAL STARTING POINT

Theorem 44. *Let $P(n)$ be a statement about a variable integer n. Suppose we fix an integer s_*. Let $\mathbb{Z}_{\geq s_*}$ denote the set of all integers n such that $n \geq s_*$. Suppose, furthermore, that*

I. $P(s_*)$ is true.

II. Any time $P(n)$ is true for one particular $n \in \mathbb{Z}_{\geq s_*}$, it follows that $P(n+1)$ is true.

Then $P(n)$ is true for every integer n belonging to $\mathbb{Z}_{\geq s_*}$.

And we can say the same thing in more formal language:

THE PRINCIPLE OF MATHEMATICAL INDUCTION WITH A GENERAL STARTING POINT (FORMAL LANGUAGE VERSION)

Theorem 44. Let $P(n)$ be a statement about a variable integer n. Suppose we fix an integer s_*. Let $\mathbb{Z}_{\geq s_*}$ denote the set of all integers n such that $n \geq s_*$. Suppose, furthermore, that

$$P(s_*) \tag{14.269}$$

and

$$(\forall n \in \mathbb{Z}_{\geq s_*})\big(P(n) \implies P(n+1)\big)\,. \tag{14.270}$$

Then

$$(\forall n \in \mathbb{Z}_{\geq s_*})P(n)\,. \tag{14.271}$$

And we can say the same thing in even more formal language:

THE PRINCIPLE OF MATHEMATICAL INDUCTION WITH A GENERAL STARTING POINT (VERY FORMAL LANGUAGE VERSION)

Theorem 44. Let $P(n)$ be a statement about a variable integer n. Let $s_* \in \mathbb{Z}$, and let

$$\mathbb{Z}_{\geq s_*} = \{n \in \mathbb{Z} : n \geq s_*\}\,. \tag{14.272}$$

Then

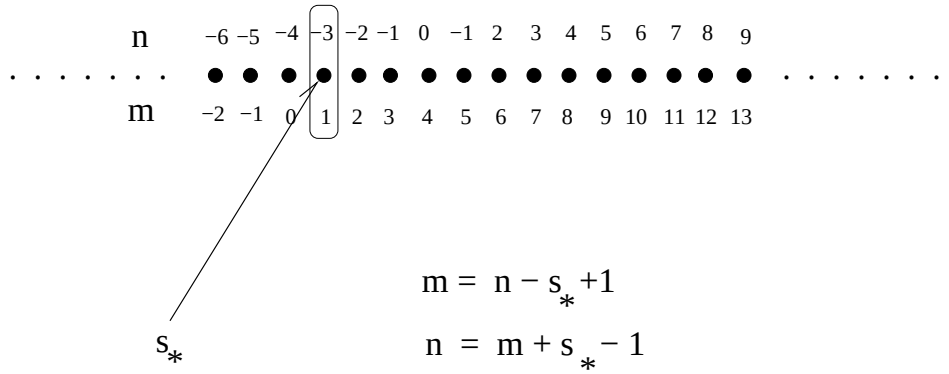
$$\Big(P(s_*) \wedge (\forall n \in \mathbb{Z}_{\geq s_*})\big(P(n) \implies P(n+1)\big)\Big) \implies (\forall n \in \mathbb{Z}_{\geq s_*})P(n)\,. \tag{14.273}$$

Proof of Theorem 44.

Assume that $P(n)$ is a 1-variable predicate and s_* is an arbitrary integer. We want to prove that if (14.269) and (14.270) hold, then (14.271) holds.

So we assume that (14.269) and (14.270) hold, and we try to prove that (14.271) holds.

We do the proof by "changing coordinates". That is, we relabel the integers so that s_* becomes 1, $s_* + 1$ becomes 2, and so on.



Changing coordinates on $\mathbb{Z}$ so that
$n = s_*$ becomes $m = 1$

(In the picture, s_* is -3.)

Precisely, we introduce a new variable m related to n by

$$m = n + 1 - s_* \,. \tag{14.274}$$

(That is: $n = s_*$ corresponds to $m = 1$, $n = s_* + 1$ corresponds to $m = 2$, and, in general, $n = s_* + k$ corresponds to $m = k$.)

We can express n in terms of m as follows:

$$n = m + s_* - 1 \,. \tag{14.275}$$

We let $Q(m)$ be $P(n)$ expressed in terms of m. That is, we let $Q(m)$ stand for $P(m + s_* - 1)$. Then $Q(1)$ is $P(s_*)$, $Q(2)$ is $P(s_* + 1)$, $Q(3)$ is $P(s_* + 2)$, and so on.

We want to prove that $P(s_*)$, $P(s_* + 1)$, $P(s_* + 2)$, ..., are all true. But this amounts to proving that $Q(1)$, $Q(2)$, $Q(3)$, ... are true, i.e. that $(\forall m \in \mathbb{N})Q(m)$.

We prove this by induction. $Q(1)$ is true because $Q(1)$ is the same as $P(s_*)$, which we are assuming is true.

And $Q(m) \implies Q(m + 1)$ is true for every $m \in \mathbb{N}$, because "$Q(m) \implies Q(m + 1)$" is equivalent to "$P(m + s_* - 1) \implies P(m + s_*)$", which is also true because $m + s_* - 1$ is to the right of s_*, so $P(m + s_* - 1)$ implies that the successor $m + s_*$ also has property P.

So $Q(m)$ satisfies all the conditions of the ordinary PMI, and we can conclude that $Q(m)$ is true for every $m \in \mathbb{N}$. And this says that $P(m+s_*-1)$ is true for all $m \in \mathbb{N}$. Hence $P(n)$ is true for all n such that $n = m + s_* - 1$ for some $m \in \mathbb{N}$. But "$n = m + s_* - 1$ for some $m \in \mathbb{N}$" is equivalent to "$n \geq s_*$"

Hence $P(n)$ is true for all $n \in \mathbb{Z}_{\geq s_*}$, and our proof is complete. **Q.E.D.**

Remark 15. Theorem 44 is a generalization of the PMI in the following precise sense: according to our definition, the set $\mathbb{Z}_{\geq 1}$ is precisely $\mathbb{N}$. So Theorem 44, if we take s_* to be 1, is exactly the PMI. □

Example 65. Let us prove the following:

Theorem 45. *If n is an integer such that $n \geq 4$, then $2^n < n!$.*

Proof. We want to prove that

$$(\forall n \in \mathbb{Z})(n \geq 4 \Longrightarrow 2^n < n!)\,. \tag{14.276}$$

Let $P(n)$ be the predicate "$2^n < n!$".

We want to prove that $(\forall n \in \mathbb{Z})(n \geq 4 \Longrightarrow P(n))$.

We are going to prove this by induction, using the PMI with a general starting point.

And we are going to take the starting point s_* to be 4.

Basis step:

We want to prove $P(4)$.
$P(4)$ says "$2^4 < 4!$".
And $2^4 = 16$, $4! = 24$, so $2^4 < 4!$.
Therefore $\boxed{P(4) \text{ is true}}$

Inductive step:

We want to prove that

$$(\forall n \in \mathbb{Z})\Big(n \geq 4 \Longrightarrow \big(P(n) \Longrightarrow P(n+1)\big)\Big)\,. \tag{14.277}$$

Let $n \in \mathbb{Z}$ be arbitrary.
We want to prove that

$$n \geq 4 \Longrightarrow (P(b) \Longrightarrow P(n+1))\,. \tag{14.278}$$

Assume that $n \geq 4$. We want to prove that $P(n) \Longrightarrow P(n+1)$.

Assume $P(n)$. We want to prove $P(n+1)$.
The inductive hypothesis $P(n)$ tells us that $2^n < n!$.
Then

$$2^{n+1} < 2n! \,. \tag{14.279}$$

But $2 \leq n+1$, so $2n! \leq (n+1)n! = (n+1)!$.
Then $2^{n+1} < (n+1)!$.
So $\boxed{P(n+1) \text{ holds}}$.

Therefore $\boxed{P(n) \Longrightarrow P(n+1)}$ (Rule $\Longrightarrow_{prove}$).

So $\boxed{n \geq 4 \Longrightarrow (P(n) \Longrightarrow P(n+1))}$ (Rule $\Longrightarrow_{prove}$).

Hence $\boxed{(\forall n \in \mathbb{Z})\Big(n \geq 4 \Longrightarrow (P(n) \Longrightarrow P(n+1))\Big)}$ (by Rule $\forall_{use}$).

This completes the inductive step.

Since we have proved that
$\boxed{P(4) \wedge (\forall n \in \mathbb{Z})\Big(n \geq 4 \Longrightarrow (P(n) \Longrightarrow P(n+1))\Big)}$,
it follows from the PMI with general starting point that $(\forall n \in \mathbb{Z})(n \geq 4 \Longrightarrow P(n))$, that is,
$\boxed{(\forall n \in \mathbb{Z})(n \geq 4 \Longrightarrow 2^n < n!)}$. **Q.E.D.**

14.2 Induction going forward and backward

The PMI says that, if a property P is true of 1, and is passed on to the right, so each natural number n passes it on to its successor $n+1$, then the property will hold of all the numbers that we reach by counting starting at 1. And the "generalized" form says that the same is true for integers if you start at any integer s_*.

It is clear that if in addition to being passed on to the right property P is also passed on to the left, (that is, if the implication $P(n+1) \Longrightarrow P(n)$ holds for every $n \in \mathbb{Z}$), then $P(n)$ will be true for every integer n.

INDUCTION GOING FORWARD AND BACKWARD

Theorem 46. *Let $P(n)$ be a statement about a variable integer n and let s_* be an integer. Suppose that*

I. $P(s_*)$ is true.

II. Any time $P(n)$ is true for one particular integer n, it follows that $P(n+1)$ is true.

III. Any time $P(n+1)$ is true for one particular integer n, it follows that $P(n)$ is true.

Then $P(n)$ is true for every integer n.

And we can say the same thing in more formal language:

INDUCTION GOING FORWARD AND BACKWARD (FORMAL LANGUAGE VERSION)

Theorem 46. Let $P(n)$ be a statement about a variable integer n and let s_* be an integer. Suppose that

$$P(s_*) \tag{14.280}$$

and

$$(\forall n \in \mathbb{Z})\big(P(n) \iff P(n+1)\big)\,. \tag{14.281}$$

Then

$$(\forall n \in \mathbb{Z})P(n)\,. \tag{14.282}$$

And we can say the same thing in even more formal language:

INDUCTION GOING FORWARD AND BACKWARD (VERY FORMAL LANGUAGE VERSION)

Theorem 46. Let $P(n)$ be a statement about a variable integer n. Let $s_* \in \mathbb{Z}$. Then

$$\Big(P(s_*) \wedge (\forall n \in \mathbb{Z})\big(P(n) \iff P(n+1)\big)\Big) \implies (\forall n \in \mathbb{Z})P(n)\,. \tag{14.283}$$

Problem 76. ***Prove*** Theorem 46. □

14.3 Examples of proofs using induction going forward and backward

14.3.1 A very simple example

Here is a simple example of a proof using indcuction going forwad and backward.

First let us review a fact that we already know:

(D3) *if $n \in \mathbb{Z}$, then $n^3 - n$ is divisible by 3.*

(This is easy to prove: we have

$$n^3 - n = n(n^2 - 1) = n(n-1)(n+1) = (n-1)n(n+1)\,,$$

so $n^3 - n$ is the product of three consecutive integers. One of these integers must be divisible by 3, so the product is divisible by 3. Actually, it is also true that $n^3 - n$ must be even, that is, divisible by 2, and then, since 2 and 3 are coprime, it follows that a stronger result is true: *$n^3 - n$ is divisible by* 6.)

In view of (D3), we may conjecture that a similar statement may be true for 4 instead of 3:

(D4) *if $n \in \mathbb{Z}$, then $n^4 - n$ is divisible by 4.*

This, however, is not true. (Proof: (D4) is a universal sentence; it says that for all integers n 4 divides $n^3 - n$. To prove that (D4) is not true, it suffices to give a counterexample. Let us just take $n = 2$. Then $2^4 = 16$, so $2^4 - 2 = 14$, which is not divisible by 4.)

How about (D5)? This one turns out to be true, and we can prove it using induction going backward and forward.

Theorem 47. *If n is an integer, then $n^5 - n$ is divisible by 5.*

Proof. We are going to use the binomial formula for the fifth power of a sum:

$$(a+b)^5 = a^5 + 5a^4b + 10a^3b^2 + 10a^2b^3 + 5ab^4 + b^5\,, \tag{14.284}$$

which is valid for all integers a, b. (And also for real numbers or, more generally, members of any commutative ring with identity.)

Using this formula we can write, for $n \in \mathbb{Z}$,

$$\begin{aligned} (n+1)^5 &= n^5 + 5n^4 + 10n^3 + 10n^2 + 5n + 1 \\ (n+1)^5 - n^5 - 1 &= 5n^4 + 10n^3 + 10n^2 + 5n \\ &= 5(n^4 + 2n^3 + 2n^2 + n)\,, \end{aligned}$$

so $(n+1)^5 - n^5 - 1$ is divisible by 5.

But

$$(n+1)^5 - (n+1) = ((n+1)^5 - n^5 - 1) + n^5 - n\,.$$

This implies that, for all $n \in \mathbb{Z}$,

$$5|(n+1)^5 - (n+1) \iff 5|n^5 - n\,. \tag{14.285}$$

In other words, the predicate "5 divides $n^5 - n$" is passed on forward (from n to $n+1$) and backward (from $n+1$ to n). This means that we are in a perfect situation to do induction going forward and backward.

Let $P(n)$ be the predicate "5 divides $n^5 - n$". We will prove the statement "$(\forall n \in \mathbb{Z})P(n)$" by induction going forward and backward. We choose the starting point s_0 to be 0.

Basis step. $P(0)$ says "5 divides 0", which is true because every integer divides 0. So $\boxed{P(0) \text{ is true}}$.

Inductive step. We have to prove that

$$(\forall n \in \mathbb{Z})\big(P(n) \iff P(n+1)\big)\,.$$

But Formula (14.285) says precisely that for every $n \in \mathbb{Z}$ $P(n) \iff P(n+1)$

This completes the inductive step. **Q.E.D.**

Problem 77. ***Prove or disprove*** each of the following statements:

1. If n is an integer, then $n^6 - n$ is divisible by 6.

2. If n is an integer, then $n^7 - n$ is divisible by 7.
3. If n is an integer, then $n^8 - n$ is divisible by 8.
4. If n is an integer, then $n^9 - n$ is divisible by 9.
5. If n is an integer, then $n^{10} - n$ is divisible by 10.
6. If n is an integer, then $n^{11} - n$ is divisible by 11.

You may find the following binomial formulas useful:

$$\begin{aligned}(a+b)^7 &= a^7 + 7a^6b + 21a^5b^2 + 35a^4b^3 + 35a^3b^4 \\ &\quad +21a^2b^5 + 7ab^6 + b^7 \\ (a+b)^{11} &= a^{11} + 11a^{10}b + 55a^9b^2 + 165a^8b^3 + 330a^7b^4 \\ &\quad +462a^6b^4 + 462a^5b^6 + 330s^4b^7 \\ &\quad +165a^3b^8 + 55a^2b^9 + 11ab^{10} + b^{11}\,.\end{aligned}$$

Remark 16. If you have done problem 77 you will have discovered the cases $p = 3, 6, 7$ and 11 of ***Fermat's little theorem:*** *If p is a prime numbet and n is an arbitrary integer then $n^p - n$ is divisible by p.* (And the case $p = 2$ is trivial, because if $n \in \mathbb{Z}$ then $n^2 - n$ is always even.) □

14.3.2 Divisibility properties of products of consecutive integers

We now discuss several theorems on divisibility of a product of consecutive integers:

1. It is easy to prove that a product $n(n+1)$ of two consecutive integers must be divisible by 2.
2. We will then look at the product $n(n+1)(n+2)$ of three consecutive integers, and prove that such a product is divisible by 6.
3. Then we will look at the product $n(n+1)(n+2)(n+3)$ of four consecutive integers, and prove that such a product is divisible by 24.
4. Since $2 = 2\times 1 = 2!$, $6 = 3\times 2\times 1 = 3!$, and $24 = 4\times 3\times 2\times 1 = 4!$, this will clearly be a good indication that there is a general pattern, namely, that for every natural number k the product of k consecutive integers is divisible by $k!$. (Recall the inductive definition of the factorial $n!$ of

a natural number: $1! = 1$ and $(n+1)! = n! \times (n+1)$ for $n \in \mathbb{N}$.) In other words, the general result should be that

$$(\forall k \in \mathbb{N})(\forall n \in \mathbb{Z})k! \,\Big|\, n(n+1)(n+2)\cdots(n+k-1) \tag{14.286}$$

or, using a notation without the mysterious and incomprehensible symbol "$\cdots$":

$$(\forall k \in \mathbb{N})(\forall n \in \mathbb{Z})k! \,\Big|\, \prod_{j=1}^{k}(n+j-1) \tag{14.287}$$

5. And we will indeed prove (14.287) eventually, but the proof will be little but harder than other proofs we have done so far, because it will use a ***double induction***: we will prove (14.287) by induction with respect to k, and for each k we will need induction with respect to n.

First let us start with the trivial result for $k = 2$:

Theorem 48. *If n is an integer, then $n(n+1)$ is even, i.e., divisible by 2. That is,*

$$(\forall n \in \mathbb{N})2|n(n+1)\,. \tag{14.288}$$

Proof. As I said earlier, this result is trivial.

Let n be an arbitrary integer.

We know that n is either even or odd.

If n is even then $\boxed{n(n+1) \text{ is even}}$.

And if n is odd then $n+1$ is even so $\boxed{n(n+1) \text{ is even}}$.

So we have proved that $n(n+1)$ is even in both cases, when n is even and when n is odd. And we know that one of these two cases must occur. So $\boxed{\boxed{n(n+1) \text{ is even}}}$.

So we have proved that $n(n+1)$ is even for an arbitrary integer n.

Hence $\boxed{\boxed{\boxed{(\forall n \in \mathbb{Z})n(n+1) \text{ is even}}}}$. **Q.E.D.**

We now want to prove that the product $n(n+1)(n+2)$ of three consecutive integers is divisible by 6. And the strategy is going to be to prove the result by induction going forward and backward.

Here is the result:

Theorem 49. *If n is an integer, then $n(n+1)(n+2)$ is divisible by 6. That is,*

$$(\forall n \in \mathbb{Z}) 6 | n(n+1)(n+2)\,. \tag{14.289}$$

Proof. Let $P(n)$ be the statement "$6|n(n+1)(n+2)$"

We prove that $(\forall n \in \mathbb{Z})P(n)$ by induction going forward and backward.

Basis step. If $n = 0$, then $n(n+1)(n+2) = 0$, so $P(0)$ is the statement "6|0", which is obviously true. So $\boxed{P(0)}$ is true.

Inductive step. We want to prove that

$$(\forall n \in \mathbb{Z})\big(P(n) \iff P(n+1)\big)\,. \tag{14.290}$$

Let n be an arbitrary integer.

We want to prove that $P(n) \iff P(n+1)$.

We already know that $n(n+1)$ is even. So we can write

$$n(n+1) = 2k\,,\ \ k \in \mathbb{Z}\,.$$

Then

$$\begin{aligned}(n+1)(n+2)(n+3) &= (n+3)(n+1)(n+2)\\ &= n(n+1)(n+2)\\ &\quad\ +3(n+1)(n+2)\\ &= n(n+1)(n+2) + 3\times 2k\\ &= n(n+1)(n+2) + 6k\,.\end{aligned}$$

If 6 divides $n(n+1)(n+2)$, then $(n+1)(n+2)(n+3)$ is the sum of two integers that are divisible by 6. So 6 divides $(n+1)(n+2)(n+3)$.

If 6 divides $(n+1)(n+2)(n+3)$, then $n(n+1)(n+2)$ is the difference of two integers that are divisible by 6. So 6 divides $n(n+1)(n+2)$.

We have shown that

$$6|(n+1)(n+2)(n+3) \iff 6|n(n+1)(n+2)\,,$$

i.e., that $P(n) \iff P(n+1)$.

Since we have shown that $P(n) \iff P(n+1)$ for an arbitrary integer n, it follows that

$$(\forall n \in \mathbb{Z})\big(P(n) \iff P(n+1)\big)\,,$$

and this completes the inductive step.

It the follows from Theorem 46 that $P(n)$ is true for all integers n. That is, (14.289) holds. **Q.E.D.**

In the proof of Theorem 49 we used the fact that if $n \in \mathbb{Z}$ then $n(n+1)$ is divisible by 2. Similarly, to prove that $(\forall n \in \mathbb{Z})24|n(n+1)(n+2)(n+3)$, the proof should use the result that $(\forall n \in \mathbb{Z})6|n(n+1)(n+2)$.

Similar results can be proved for the products of four and five consecutive integers.

Theorem 50. *If n is an integer, then the product $n(n+1)(n+2)(n+3)$ is divisible by 24. That is,*

$$(\forall n \in \mathbb{Z})24|n(n+1)(n+2)(n+3)\,. \tag{14.291}$$

Proof. **YOU DO THIS ONE.**

In the proof of Theorem 49 we used the fact that if $n \in \mathbb{Z}$ then $n(n+1)$ is divisible by 2.

Similarly, to prove that

$$(\forall n \in \mathbb{Z})24|n(n+1)(n+2)(n+3)\,,$$

the proof should use the result of Theorem 49, that is, that $(\forall n \in \mathbb{Z})6|n(n+1)(n+2)$.

Problem 78. ***Prove*** Theorem 50. You are ***not*** allowed to use Theorem 52.

NOTE: Theorem 50 is a special case of Theorem 52, for $k = 4$. But I want you to prove Theorem 50 directly, without using Theorem 52. □

Theorem 51. *If n is an integer, then the product $n(n+1)(n+2)(n+3)(n+4)$ is divisible by 120. That is,*

$$(\forall n \in \mathbb{Z})120|n(n+1)(n+2)(n+3)(n+4)\,. \tag{14.292}$$

Proof. **YOU DO THIS ONE.**

In the proof of Theorem 50 we used the fact that if $n \in \mathbb{Z}$ then $n(n+1)(n+2)$ is divisible by 6. Similarly, to prove that $(\forall n \in \mathbb{Z})120|n(n+1)(n+2)(n+3)(n+4)$, the proof should use the result that

$$(\forall n \in \mathbb{Z})24|n(n+1)(n+2)\,.$$

Problem 79. ***Prove*** Theorem 51. You are ***not*** allowed to use Theorem 52.

NOTE: Theorem 51 is a special case of Theorem 52, for $k = 5$. But I want you to prove Theorem 51 directly, without using Theorem 52. □

What we have done so far is clearly the beginning of a proof by induction. We have proved the following:

(*) *for $k = 1, 2, 3, 4, 5$ the product of k consecurive integers is divisible by $k!$.*

This makes it natural to make the following

Conjecture. *For every natural number k the product of k consecutive integers is divisible by $k!$.*

But, of course, knowing that something is true for a few values of k in no way proves that it is true for all k, If we want to be sure that a statement about k is true for all k, we have to prove it.

So let us prove it.

Theorem 52. *If k is a natural number then every product of k consecutive integers is divisible by $k!$.*

Proof. As usual, our first task is to rewrite the statement we want to prove in precise formal language. And for that purpose we need to write a formula for the product of k consecutive integers.

If we start with an integer n, then the k consecutive integers starting at n are n, $n+1$, $n+2$, ..., up to $n+k-1$. And the product of these k integers is $\prod_{j=1}^{k}(n+j-1)$. (For example, for $k = 3$, the product is $n(n+1)(n+2)$. The first factor is n, that is $n+j-1$ with $j = 1$, and the last factor is $n+2$, that is, $n+j-1$ with $j = 3$.)

Let us call this product $a_{n,k}$, so

$$a_{n,k} = \prod_{j=1}^{k} (n + j - 1)\,, \tag{14.293}$$

or, if you prefer,

$$a_{n,k} = n \times (n+1) \times (n+2) \times \cdots \times (n+k-1)\,. \tag{14.294}$$

So, for example,

$$\begin{aligned} a_{2,3} &= 2 \times 3 \times 4\,, \\ a_{-5,7} &= (-5) \times (-4) \times (-3) \times (-2) \times (-1) \times 0 \times 1\,, \\ a_{4,9} &= 4 \times 5 \times 6 \times 7 \times 8 \times 9 \times 10 \times 11 \times 12\,. \end{aligned}$$

Then what we want to prove is the following statement:

$$(\forall k \in \mathbb{N})(\forall n \in \mathbb{Z})\; k! \,\Big|\, a_{n,k}\,. \tag{14.295}$$

In order to prove this, we will use induction.

We let $P(k)$ be the predicate "for every integer n, the product of k consecutive integers starting with n is divisible by $k!$". That, $P(k)$ is the predicate

$$(\forall n \in \mathbb{Z})\; k! \,\Big|\, a_{n,k}\,. \tag{14.296}$$

Basis step of the induction. We want to prove that $P(1)$ is true. And $P(1)$ is true, for trivial reasons: $P(1)$ says "$(\forall n \in \mathbb{Z})1! \,\Big|\, a_{n,1}$", i.e., "$(\forall n \in \mathbb{Z})1 \,\Big|\, n$", and this is true because every integer is divisible by 1. So we have proved $\boxed{P(1)}$.

Inductive step. We want to prove that

$$(\forall k \in \mathbb{N})\big(P(k) \Longrightarrow P(k+1)\big)\,. \tag{14.297}$$

Let $k \in \mathbb{N}$ be arbitrary. We want to prove that

$$P(k) \Longrightarrow P(k+1)\,. \tag{14.298}$$

Assume $P(k)$. That is, we assume that the product of k consecutive integers is divisible by $k!$.

We want to prove $P(k+1)$. That is, we want to prove

$$(\forall n \in \mathbb{Z})\,(k+1)! \,\Big|\, a_{n,k+1}\,. \tag{14.299}$$

We are going to prove this by induction going forward and backward. This means that

- We are going to do a second induction proof, with respect to n, within the main proof by induction with respect to to k.
- We are going to call this "the n-induction", to distinguish it from the main induction, the "k-induction".

So at this point

- we are within the k-induction,
- we are about to do the n-induction,
- we are assuming that $P(k)$ is true,
- and we are trying to prove that $P(k+1)$ is true, that is, we are trying to prove that (14.299) is true,
- and, since (14.299) is a universal sentence about "all integers n", we are going to do the proof by induction going forward and backward.

We let $Q(n)$ be the predicate

$$(k+1)! \,\Big|\, a_{n,k}\,. \tag{14.300}$$

We choose the starting point s_* of our induction to be 0.

Basis step of the n-induction. We want to prove that $Q(0)$ is true. But $Q(0)$ says

$$(k+1)! \,\Big|\, a_{0,k+1}\,.$$

And $a_{0,k+1} = 0$, because $a_{0,k+1}$ is a product of numbers the first one of which is 0. So $Q(0)$ says "$(k+1)!\Big|0$", and this is true, because 0 is divisible by every integer. So we have proved $\boxed{Q(0)"}$.

Inductive step of the n-induction. We want to prove that

$$(\forall n \in \mathbb{Z})\big(Q(n) \iff Q(n+1)\big)\,. \tag{14.301}$$

Let $n \in \mathbb{Z}$ be arbitrary. We want to prove

$$Q(n) \iff Q(n+1)\,. \tag{14.302}$$

$Q(n)$ says that $(k+1)!$ divides $a_{n,k+1}$.
And $Q(n+1)$ says that $(k+1)!$ divides $a_{n+1,k+1}$.
We are going to prove that

$$(k+1)! \text{ divides } a_{n+1,k+1} - a_{n,k+1}\,. \tag{14.303}$$

Before we do that, let me explain why this is a significant fact.
Suppose that we have proved (14.303).
We are going to prove the two implicatios $Q(n) \Longrightarrow Q(n+1)$ and $Q(n+1) \Longrightarrow Q(n)$.

First, assume that $Q(n)$ holds.

Then $a_{n,k+1}$ is divisible by $(k+1)!$.

Since $a_{n+1,k+1} - a_{n,k+1}$ is also divisible by $(k+1)!$, we can conclude that the sum $a_{n,k+1} + \big(a_{n+1,k+1} - a_{n,k+1}\big)$ is divisible by $(k+1)!$.

But this sum is equal to $a_{n+1,k+1}$, So $a_{n+1,k+1}$ is divisible by $(k+1)!$.

That says that $Q(n+1)$ holds.

Hence $Q(n) \Longrightarrow Q(n+1)$.

Conversely, assume $Q(n+1)$ holds. Then $a_{n+1,k+1}$ is divisible by $(k+1)!$.

Since the difference $a_{n+1,k+1} - a_{n,k+1}$ is divisible by $(k+1)!$, we can conclude that $a_{n+1,k+1} - \big(a_{n+1,k+1} - a_{n,k+1}\big)$ is divisible by $(k+1)!$.

But

$$a_{n+1,k+1} - \big(a_{n+1,k+1} - a_{n,k+1}\big) = a_{n,k+1}\,.$$

So $a_{n,k+1}$ is divisible by $(k+1)!$.

That says that $Q(n)$ holds.

So $Q(n+1) \Longrightarrow Q(n)$.
Summarizing, we have shown that, if the assertion (14.303) is true, then both implications "$Q(n) \Longrightarrow Q(n+1)$" and "$Q(n+1) \Longrightarrow Q(n)$" hold, so $Q(n) \Longleftrightarrow Q(n+1)$, which is exactly what we are trying to prove to complete the n-induction.
In other words: *all we need to do is prove (14.303) and that will complete our proof.*

We now prove (14.303).
The number $a_{n,k+1}$ is the product of $k+1$ consecutive integers starting with n and ending with $n+k$. That is,

$$a_{n,k+1} = n \times (n+1) \times (n+2) \times \cdots \times (n+k-1) \times (n+k)\,.$$

And then

$$a_{n,k+1} = n \times \Big((n+1) \times (n+2) \times \cdots \times (n+k-1) \times (n+k)\Big)\,,$$

so $a_{n,k+1}$ is equal to n times the product $(n+1)\times(n+2)\times\cdots\times(n+k-1)\times(n+k)$ of k consecutive integers starting with $n+1$. That is,

$$a_{n,k+1} = n \times a_{n+1,k}\,. \tag{14.304}$$

Similarly, the number $a_{n+1,k+1}$ is the product of $k+1$ consecutive integers starting with $n+1$ and ending with $n+k+1$. That is,

$$a_{n+1,k+1} = (n+1) \times (n+2) \times \cdots \times (n+k) \times (n+k+1)\,.$$

So

$$a_{n+1,k+1} = \Big((n+1) \times (n+2) \times \cdots \times (n+k)\Big) \times (n+k+1)\,.$$

In other words, $a_{n+1,k+1}$ is equal to the product of k consecutive integers starting with $n+1$, multiplied by $n+k+1$. That is,

$$a_{n+1,k+1} = a_{n+1,k} \times (n+1+k)\,. \tag{14.305}$$

Therefore

$$\begin{aligned} & a_{n+1,k+1} - a_{n,k+1} \\ = \;& a_{n+1,k} \times (n+1+k) - n \times a_{n+1,k} \\ = \;& (n+k+1) \times a_{n+1,k} - n \times a_{n+1,k} \\ = \;& \big((n+k+1) - n\big) \times a_{n+1,k} \\ = \;& (k+1) \times a_{n+1,k}\,. \end{aligned}$$

So we get the key formula

$$a_{n+1,k+1} - a_{n,k+1} = (k+1) \times a_{n+1,k}\,. \tag{14.306}$$

(see the example in the box below to get a better understanding of this formula).

THE FORMULA
$a_{n+1,k+1} - a_{,k+1} = (k+1) \times a_{n+1,k}$:
AN EXAMPLE

Take $n = 11$, $k = 5$. Then

$$\begin{aligned}
a_{11,6} &= 11 \times 12 \times 13 \times 12 \times 15 \times 16\,, \\
a_{12,6} &= 12 \times 13 \times 12 \times 15 \times 16 \times 17\,, \\
a_{11,6} &= 11 \times \Big(12 \times 13 \times 12 \times 15 \times 16\Big) \\
&= 11 \times a_{12,5} \\
a_{12,6} &= \Big(12 \times 13 \times 12 \times 15 \times 16\Big) \times 17 \\
&= 17 \times \Big(12 \times 13 \times 12 \times 15 \times 16\Big) \\
&= 17 \times a_{12,5}\,,
\end{aligned}$$

so

$$\begin{aligned}
a_{12,6} - a_{11,6} &= (17 - 11) \times a_{12,5} \\
&= 6 \times a_{12,5}\,.
\end{aligned}$$

That is,

$$\boxed{a_{n+1,k+1} - a_{n,k+1} = (k+1) \times a_{n+1,k}}.$$

Now comes ***the crucial point of the proof:*** remember that we are within the k-induction. We are assuming $P(k)$ and trying to prove $P(k+1)$. So at this point we are allowed to use $P(k)$. And $P(k)$ says that

$$(\forall n \in \mathbb{Z})\, k! \Big| a_{n,k}\,. \tag{14.307}$$

So we can use (14.307).
Then $k!$ divides $a_{n+1,k}$, so we can write

$$a_{n,k} = m \times k!$$

for some $m \in \mathbb{Z}$. Then

$$\begin{aligned}
a_{n+1,k+1} - a_{n,k+1} &= (k+1) \times k! \times m \\
&= (k+1)! \times m\,,
\end{aligned}$$

so $(k+1)!$ divides $a_{n+1,k+1} - a_{n,k+1}$.

That is, we have proved (14.303) and, as was explained before, it follows from this that

$\boxed{Q(n) \iff Q(n+1)}$.

Since we have proved that $Q(n) \iff Q(n+1)$ for an arbitrary integer n, we can conclude that $\boxed{(\forall n \in \mathbb{Z})\big(Q(n) \iff Q(n+1)\big)}$.

This completes the inductive step of the n-induction.

We have proved that $Q(0)$ and also that $(\forall n \in \mathbb{Z})\big(Q(n) \iff Q(n+1)\big)$. By the PMI Going Forward and Backward, it follows that

$$(\forall n \in \mathbb{Z})Q(n)\,. \tag{14.308}$$

Since $Q(n)$ is the predicate "$(k+1)!\Big|a_{n,k+1}$", we have proved

$$(\forall n \in \mathbb{Z})(k+1)!\Big|a_{n,k+1}\,, \tag{14.309}$$

that is, we have proved $P(k+1)$.

Since we have proved $P(k+1)$ assuming $P(k)$, it follows that

$$P(k) \Longrightarrow P(k+1)\,. \tag{14.310}$$

Since we have proved (14.310) for an arbitrary natural numebr k, it follows that

$$(\forall k \in \mathbb{N})\big(P(k) \Longrightarrow P(k+1)\big)\,. \tag{14.311}$$

So we have proved $P(1)$, and we have also proved that $(\forall k \in \mathbb{N})\big(P(k) \Longrightarrow P(k+1)\big)$. It follows from the PMI that

$$(\forall k \in \mathbb{N})P(k)\,. \tag{14.312}$$

But $P(k)$ is the predicate "$(\forall n \in \mathbb{Z})\, k!\Big|a_{n,k}$".

So we have proved

$$(\forall k \in \mathbb{N})(\forall n \in \mathbb{Z})\, k!\Big|a_{n,k}\,, \tag{14.313}$$

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

14.4 An application of Theorem 52: integrality of the binomial coefficients

An important application of Theorem 52, on the divisibility of a product of k consecutive integers, is to give a second proof of Theorem 54, different from the one suggested in the hints for Problem 80.

14.4.1 The binomial coefficients

The binomial coefficients $\binom{n}{k}$ are defined as follows:

Definition 22. *If n, k are nonnegative integers*[83] *such that $k \leq n$, then the binomial coefficient $\left(\begin{array}{c} n \\ k \end{array} \right)$ is defined by the formula*

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}. \tag{14.314}$$

Remark 17. One of the most important facts about the numbers $\left(\begin{array}{c} n \\ k \end{array} \right)$ is that ***they are always integers***.

It is not obvious at all from Definition 22 that $\binom{n}{k}$ is always an integer.

For example: ***why should*** $\binom{17}{9}$ ***be an integer? Why does*** $17!$ ***have to be divisible by*** $9! \times 8!$***?*** There is no doubt that $17!$ has to be divisible by $9!$, because $17! = 17 \times 16 \times 15 \times 14 \times 13 \times 12 \times 11 \times 10 \times 9!$. But why is the quotient

$$\frac{17!}{9!} = 17 \times 16 \times 15 \times 14 \times 13 \times 12 \times 11 \times 10$$

divisible by $8!$? In this particular example, it is easy to do the cancellations,

[83] A nonnegative integer is an integer n such that $n \geq 0$. So the nonnegative integers are the natural numbers, together with the integer 0, which is not a natural number. The set of all nonnegative integers is denoted by the expression "$\mathbb{N} \cup \{0\}$". Therefore "$n \in \mathbb{N} \cup \{0\}$" is a way of saying that $n \in \mathbb{N} \vee n = 0$, i.e., that n is a nonnegative integer.

and get

$$\begin{aligned}
\frac{17!}{8!9!} &= \frac{17\times16\times15\times14\times13\times12\times11\times10}{8\times7\times6\times5\times4\times3\times2} \\
&= \frac{17\times15\times14\times13\times12\times11\times10}{7\times6\times5\times4\times3} \\
&= \frac{17\times14\times13\times12\times11\times10}{7\times6\times4} \\
&= \frac{17\times14\times13\times12\times11\times5}{7\times6\times2} \\
&= \frac{17\times13\times12\times11\times5}{6} \\
&= 17\times13\times2\times11\times5\,.
\end{aligned}$$

So in this particular case it is clear that $\left(\begin{array}{c} 17 \\ 9 \end{array}\right)$ is an integer, but ***it is not clear yet why it should be true in general that*** $\left(\begin{array}{c} n \\ k \end{array}\right)$ ***is an integer for all*** $n, k \in \mathbb{N}\cup\{0\}$ ***such that*** $k \le n$.

The following two theorems give one answer to this question. □

Theorem 53. *Let* $n, k \in \mathbb{N}\cup\{0\}$ *be such that* $1 \le k \le n$. *Then*

$$\left(\begin{array}{c} n+1 \\ k \end{array}\right) = \left(\begin{array}{c} n \\ k-1 \end{array}\right) + \left(\begin{array}{c} n \\ k \end{array}\right). \tag{14.315}$$

Proof. **YOU DO IT.**

Theorem 54. *If* n, k *are nonnegative integers such that* $k \le n$, *then the binomial coefficient* $\left(\begin{array}{c} n \\ k \end{array}\right)$ *is an integer.*

Proof. **YOU DO IT.**

Problem 80. ***Prove*** Theorems 53 and 54.

The proof of Theorem 53 should be very easy: you just add the fractions $\frac{n!}{(k-1)!(n-(k-1))!}$ and $\frac{n!}{k!(n-k)!}$ and the answer turns out to be $\frac{(n+1)!}{k!(n-k)!}$. ***This is <u>not</u> a proof by induction.***

The proof of Theorem 54 should be very easy, by induction. Theorem 53 easily implies that if all the binomial coefficients $\left(\begin{array}{c} n \\ k \end{array}\right)$ are integers for

a given n, then all the binomial coefficients $\left(\begin{array}{c} n+1 \\ k \end{array} \right)$ are integers as well. And this is basically the inductive step.

But ***you should write the proof carefully and correctly.*** In particular, pay attention to the fact that what you want to prove is a statement with ***two quantifiers***, but in a proof by induction of $(\forall n \in \mathbb{N} \cup \{0\})P(n)$, the sentence $P(n)$ has to have n as an open variable, and no other open variables. So you cannot take $P(n)$ to be a closed formula such as

$$(\forall n \in \mathbb{N} \cup \{0\})(\forall j \in \mathbb{N} \cup \{0\})\Big(k \leq n \Longrightarrow \left(\begin{array}{c} n \\ k \end{array} \right) \in \mathbb{Z}\Big)\,,$$

and you cannot take $P(n)$ to be "$k \leq n \Longrightarrow \left(\begin{array}{c} n \\ k \end{array} \right) \in \mathbb{Z}$" either, because this formula has two open variables.

Also, you should pay attention in your inductive step to the fact that Formula (14.315) cannot be applied if $k = 0$, so you will have to consider the case when $k = 0$ separately. □

14.4.2 A second proof of the integrality of the binomial coefficients

We want to prove that the binomial coefficients $\left(\begin{array}{c} n \\ k \end{array} \right)$ are integers, for $n, k \in \mathbb{N} \cup \{0\}$ and $k \leq n$.

First we write

$$\begin{aligned} n! &= 1 \times 2 \times \cdots \times (n-k) \times (n+1-k) \times \cdots \times n \\ &= \big(1 \times 2 \times \cdots \times (n-k)\big) \times \big((n+1-k) \times \cdots \times n\big) \\ &= (n-k)! \times \big((n+1-k) \times \cdots \times n\big)\,. \end{aligned}$$

We then observe that $(n+1-k) \times \cdots \times n$ is the product of k consecutive integers starting at $n+1-k$, which is the number that in the proof of Theorem 52 we called $a_{n+1-k,k}$.

In other words,

$$n! = (n-k)! \times a_{n+1-k,k}\,. \tag{14.316}$$

Finally, we use Theorem 52 to conclude that $a_{n+1-k,k}$ is divisible by $k!$. Hence we can write

$$a_{n+1-k,k} = k! \times m\,,$$

where m is an integer.

It then follows that

$$\begin{aligned} n! &= (n-k)! \times k! \times m \\ &= \big((n-k)! \times k!\big) \times m\,. \end{aligned}$$

so $n!$ is divisible by $(n-k)! \times k!$, and this completes the proof of Theorem 54. **Q.E.D.**

14.5 Strong induction (a.k.a. "complete induction")

Suppose we are trying to prove a proposition that is of the form $(\forall n \in \mathbb{N})P(n)$. It may happen that we cannot prove the implication $P(n) \Longrightarrow P(n+1)$, because property P is not inherited by $n+1$ from n for every n, but the property is inherited by $n+1$ from some previous natural number, such as $n-1$, or $n-2$. Then ***it still follows that*** $(\forall n \in \mathbb{N})P(n)$.

Example 66. Let $P(n)$ be the predicate[84] "$n = 1 \vee n$ is a product of prime numbers".

We would like to prove that

$$(\forall n \in \mathbb{N})P(n)\,, \tag{14.317}$$

that is, that

$$(\forall n \in \mathbb{N})\big(n = 1 \vee n \text{ is a product of prime numbers}\,,$$

or, equivalently,

$$(\forall n \in \mathbb{N})\Big(n \geq 2 \Longrightarrow \big(n \text{ is a product of prime numbers}\big)\Big)\,.$$

(That is, "if n is a natural number and $n \geq 2$ then n is a product of prime numbers.)

To prove this, we would like to use induction. The basis step is easy: $P(1)$ is true, because $P(1)$ says "$1 = 1 \vee 1$ is a product of prime numbers", and this is obviously true because $1 = 1$.

But when we get to the inductive step, and we try to prove that implication $P(n) \Longrightarrow P(n+1)$ for every n, we get into trouble.

Look, for example, at $n = 47$ and $n = 60$. We want to prove that $P(47) \Longrightarrow P(48)$ and $P(60) \Longrightarrow P(61)$. But, although $P(48)$ and $P(61)$ are true (because $48 = 2 \times 2 \times 2 \times 3$, and 61 is prime, so both 48 and 61 are products of primes), the reasons that $P(48)$ and $P(61)$ are true have nothing to do with the facts that $P(47)$ and $P(60)$ are true.

[84] The precise meaning of "is a product of prime numbers" was defined in section 2.3.3, Definition 6, on page 22. In particular, we insist on the fact that ***a single prime number is a product of primes according to our definition.***

Indeed:

- $P(48)$ is true because
  - $48 = 8 \times 6$,
  - 8 and 6, are both products of primes, ***because all the natural numbers that are*** ≤ 47 ***are products of primes***,
  - so 48 is a product of primes.
- And $P(61)$ is true because
  - 61 is prime. □

So, as Example 66 shows, it is not going to be possible to prove the implication $P(n) \Longrightarrow P(n+1)$ for every n.

On the other hand, if we associate to the predicate $P(n)$ another predicate, $Q(n)$, defined by

$$Q(n) \text{ means } \text{“}P(1) \wedge P(2) \wedge \cdots \wedge P(n)\text{”} ,$$

that is,

$$Q(n) \text{ means } \text{“}(\forall k \in \mathbb{N})\big(k \leq n \Longrightarrow P(k)\big)\text{”}.$$

then it is clear that

$\left(*\right)$ ***if we prove that*** $(\forall n \in \mathbb{N})Q(n)$***, then it follows that*** $(\forall n \in \mathbb{N})P(n)$***.***

(Why? Suppose that $(\forall n \in \mathbb{N})Q(n)$. Let $n \in \mathbb{N}$ be arbitrary. Then $Q(n)$ is true, by Rule $\forall_{use}$. Therefore the proposition $P(1) \wedge P(2) \wedge \cdots \wedge P(n)$ is true, so in particular $P(n)$ is true. Hence $(\forall n \in \mathbb{N})P(n)$.)

Furthermore,

$\left(**\right)$ ***To prove that*** $(\forall n \in \mathbb{N})Q(n)$ ***by induction, in the inductive step, when we want to prove that the implication*** $Q(n) \Longrightarrow Q(n+1)$ ***is true, it suffices to prove that the weaker implication*** $Q(n) \Longrightarrow P(n+1)$ ***is true.***

(Why? Let us assume that $Q(n)$. We want to prove that $Q(n+1)$. That is, we need to prove the conjunction “$P(1) \wedge P(2) \wedge \cdots \wedge P(n) \wedge P(n+1)$”. But we already know that “$P(1) \wedge P(2) \wedge \cdots \wedge P(n)$” is true, because that is what $Q(n)$ is. So all we need in order to prove $Q(n+1)$ is to prove $P(n+1)$.)

Strong Induction (a.k.a. "complete induction")

Let $P(n)$ be a one-variable predicate.
Let $Q(n)$ be the predicate

$$P(1) \wedge P(2) \wedge \cdots \wedge P(n)\,,$$

so that $Q(n)$ means

$$(\forall k \in \mathbb{N})\big(k \leq n \Longrightarrow P(k)\big)\,.$$

Then, if

$$P(1)$$

and

$$(\forall n \in \mathbb{N})\big(Q(n) \Longrightarrow P(n+1)\big)\,,$$

it follows that $(\forall n \in \mathbb{N})P(n)$.

Example 67. Let us prove

Theorem 55. *If n is a natural number and $n \geq 2$ then n is a product of prime numbers.*

Proof. Let $P(n)$ be the predicate "if $n \geq 2$ then n is a product of prime numbers".

Let $Q(n)$ be the predicate "$P(k)$ is true for all natural numbers k such that $k \leq n$".

We prove $(\forall n \in \mathbb{N})P(n)$ using strong induction.

For this purpose, we prove the two propositions $P(1)$ and $(\forall n \in \mathbb{N})\big(Q(n) \Longrightarrow P(n+1)\big)$.

Basis step. We have to prove $P(1)$. But $P(1)$ says "if $1 \geq 2$ then 1 is a product of prime numbers", and this is an implication with a false premise. So $P(1)$ is true.

Inductive step. We have to prove that

$$(\forall n \in \mathbb{N})\big(Q(n) \Longrightarrow P(n+1)\big)\,. \tag{14.318}$$

Let $n \in \mathbb{N}$ be arbitrary. We want to prove that $Q(n) \Longrightarrow P(n+1)$.

Assume $Q(n)$. We want to prove $P(n+1)$.

So we want to prove that $n+1$ is a product of prime numbers.

But $n+1$ is either prime, or not prime.

If $n+1$ is prime, then it is a product of primes, and $P(n+1)$ holds.

If $n+1$ is not prime, then, since $n+1 \neq 1$, it follows that $n+1$ is the product $j \times k$ of two natural numbers that are both > 1.

Clearly, then, $j \leq n$ and $k \leq n$. (If j was $> n$, then j would be $\geq n+1$ and, since $k > 1$, it would follow that $jk > n+1$. But this is not possible, because $jk = n+1$. So $j \leq n$. A similar argument proves that $k \leq n$.)

Since $Q(n)$ holds, both j and k are products of primes.

And then $n+1$, the product of j and k, is also a product of primes.

So $P(n+1)$ holds.

We have proved that $P(n+1)$ holds in both cases, when $n+1$ is prime and when $n+1$ is not prime.

Hence we have proved $P(n+1)$, assuming $Q(n)$.

So we have proved that $Q(n) \implies P(n+1)$, assuming that n is an arbitrary natural number.

Hence we have proved $(\forall n \in \mathbb{N})\big(Q(n) \implies P(n+1)\big)$, completing the inductive step.

Since we have proved both $P(1)$ and $(\forall n \in \mathbb{N})\big(Q(n) \implies P(n+1)\big)$, it follows from the strong principle of mathemtical induction that $(\forall n \in \mathbb{N})P(n)$, that is,

$$(\forall n \in \mathbb{N})\big(n \geq 2 \implies n\,, \text{is a product of primes}\,.$$

This completes our proof. **Q.E.D.**

14.5.1 Stronger and weaker statements

Remark 18. Why did I say that the implication $Q(n) \implies P(n+1)$ is "weaker" than the implication $P(n) \implies P(n+1)$?

Intuitively, a proposition A is weaker than a proposition B if it gives less information. This means that knowing that B is true tells us that A is true, so if we know that B is true then we know that A is true. (So if we know B

then we know B and A, but if we know A we only know A; we don't know B.)

More formally, we have

Definition 23. *A proposition A is weaker than a proposition B if the proposition $B \Longrightarrow A$ is true. And in that case we also say that B is stronger than A.* □

Example 68. Let A be the proposition "you got a passing grade", let B be the proposition "you got an 'A' grade". Which one gives you more information? Obviously, B does. So A should be weaker than B, and B should be stronger than A.

And, indeed, the proposition $B \Longrightarrow A$ is clearly true. So A is weaker than B according to our definition. □

Returning now to $P(n)$ and $Q(n)$, it is clear that

$$\Big(P(n) \Longrightarrow P(n+1)\Big) \Longrightarrow \Big(Q(n) \Longrightarrow P(n+1)\Big). \tag{14.319}$$

(Proof: Assume that $P(n) \Longrightarrow P(n+1)$. We want to prove that $Q(n) \Longrightarrow P(n+1)$. Assume $Q(n)$. We want to prove $P(n+1)$. Clearly, $Q(n) \Longrightarrow P(n)$. Since we are assuming $Q(n)$, it follows from the Modus Ponens rule—i.e., Rule $\Longrightarrow_{use}$—that $P(n)$ is true. Since we are assuming that $P(n) \Longrightarrow P(n+1)$, it follows again from the Modus Ponens rule that $P(n+1)$. So we have proved $Q(n) \Longrightarrow P(n+1)$, assuming $P(n) \Longrightarrow P(n+1)$. Hence (14.319) holds.)

So we see that "$Q(n) \Longrightarrow P(n+1)$" is weaker than "$P(n) \Longrightarrow P(n+1)$" in the very precise sense of Definition 23. □

Problem 81. For each of the following pairs A, B of propositions, indicate which one is stronger and which one is weaker. (You may assume that n and f are arbitrary objects that have been given to you, that is, they are fixed objects but you do not know who they are.)

1. A is "n is a natural number" and B is "n is an integer".
2. A is "if n is a natural number then $n > 0$" and B is "if n is an integer then $n > 0$".
3. A is "f is a continuous function on an interval $[a, b]$" and B is "f is a differentiable function on an interval $[a, b]$".

4. A is "every continuous function on an interval $[a, b]$ has a maximum and a minimum on $[a, b]$", and B is "every differentiable function on an interval $[a, b]$ has a maximum and a minimum on $[a, b]$". □

Problem 82. ***Prove, using the 14 rules of logic***, that

1. *If A, B, C are propositions, then if A is weaker than B then $A \Longrightarrow C$ is stronger than $B \Longrightarrow C$.* (See also Example 69 below.)

2. *If A, B, C are propositions, then if B is stronger than C it follows that $A \Longrightarrow B$ is stronger than $A \Longrightarrow C$.*

3. *If A, B, C, D are propositions, then if B is stronger than A and C is stronger than D it follows that $A \Longrightarrow C$ is stronger than $B \Longrightarrow D$.*

4. *If A, B, C are propositions, then if A is weaker than B then $A \wedge C$ is weaker than $B \wedge C$.*

5. *If $X(n)$ and $Y(n)$ are predicates with the open variable n* (so that for each fixed n $X(n)$ and $Y(n)$ are propositions) then if $X(n)$ is weaker that $Y(n)$ for each n in some set S, *it follows that the proposition "$(\forall n \in S)X(n)$" is weaker than "$(\forall n \in S)Y(n)$" and the proposition "$(\exists n \in S)X(n)$" is weaker than "$(\exists n \in S)Y(n)$".* □

Example 69. Why is strong induction called "strong induction"?

The reason is this:

- Clearly, for each $n \in \mathbb{N}$ the proposition $Q(n)$ is stronger than $P(n)$.
- Hence for each $n \in \mathbb{N}$ the implication "$Q(n) \Longrightarrow P(n+1)$" is weaker than "$P(n) \Longrightarrow P(n+1)$" (because of the first result of Problem 82).
- So "$(\forall n \in \mathbb{N})\big(Q(n) \Longrightarrow P(n+1)\big)$" is weaker than "$(\forall n \in \mathbb{N})\big(P(n) \Longrightarrow P(n+1)\big)$" (because of the third result of Problem 82).
- Hence "$P(1) \wedge (\forall n \in \mathbb{N})\big(Q(n) \Longrightarrow P(n+1)\big)$" is weaker than "$P(1) \wedge (\forall n \in \mathbb{N})\big(P(n) \Longrightarrow P(n+1)\big)$" (because of the second result of Problem 82).
- And then the implication

$$\Big(P(1) \wedge (\forall n \in \mathbb{N})\big(Q(n) \Longrightarrow P(n+1)\big)\Big) \Longrightarrow (\forall n \in \mathbb{N})P(n) \tag{14.320}$$

is stronger than the implication

$$\Big(P(1) \wedge (\forall n \in \mathbb{N})\big(P(n) \Longrightarrow P(n+1)\big)\Big) \\ \Longrightarrow (\forall n \in \mathbb{N})P(n)\,. \tag{14.321}$$

But (14.321) is the ordinary Principle of Mathematical Induction, and (14.320) is the strong Principle of Mathematical Induction.

So the strong PMI is indeed stronger than the ordinary PMI. □

15 The main theorems of elementary integer arithmetic I: the division theorem

We now study the phenomena that make the natural numbers and the integers different in crucial ways from the real numbers. The root of this difference is that the division operation on $\mathbb{N}$ and $\mathbb{Z}$ is very different from division on $\mathbb{R}$.

15.1 What is the division theorem about?

The first important fact about the integers is the ***division theorem***. It deals with an issue that you know very well, namely, what happens if you have an integer a and an integer b and you want to "divide" a by b:

1. First of all: dividing by zero is never a good idea, so we have to work with integers a and b such that $b \neq 0$.

2. Dividing a by b should amount, roughly, to finding a number q, called the "quotient of a by b", such that

$$a = bq\,. \tag{15.322}$$

3. If we were dealing with real numbers rather than integers, then it is always possible[85] to find q. The real number q that satisfies (15.322) is denoted by the expression $\frac{a}{b}$, that we read as "a over b", or "a divided by b".

4. The situation is different when we are dealing with integers rather than real numbers. In this case, it is not always possible to find an integer q for which (15.322) is satisfied *exactly*. But we can come close: we can find an integer q for which (15.322) is satisfied *approximately*.

5. Precisely, let us rewrite (15.322) as follows:

$$a = bq + r \qquad \text{and} \quad r = 0\,. \tag{15.323}$$

   Then what happens is this: we cannot satisfy (15.323), but we can satisfy

$$a = bq + r \qquad \text{and} \quad r \text{ is small}\,. \tag{15.324}$$

[85] Assuming, of course, that $b \neq 0$.

6. And the precise meaning of "small", if $b > 0$, is "$0 \leq r < b$". So what you will be satisfying (if $b > 0$) is

$$a = bq + r \qquad \text{and} \quad 0 \leq r < b\,. \tag{15.325}$$

7. The number q is called the ***quotient of the division of*** a ***by*** b, and the number r is called the ***remainder of the division of*** a ***by*** b.

8. The reason that r is called the "remainder" is very straightforward: suppose you have, say, 27 dollar bills, and you want to divide them equally among 5 people. Then the best you can do is give 5 dollars to each of the five people, and when you do that 2 dollars will "remain".

9. Notice that, if instead of 27 dollar bills you were dealing with, say, 27 gallons of water, then you would be able to divide the water equally, by giving 5.4 gallons to each of the five people. But with dollar bills you cannot do that. That's because ***dollar bills are countable***, whereas ***water is uncountable***. In other words,

   - You can talk about the ***amount*** of water in a tank, and ***amounts of water are measured in terms of real numbers***.
   - And you cannot talk about the ***number*** of water in a tank.
   - You can talk about the ***number*** of dollar bills in your wallet, and ***numbers of dollar bills are measured in terms of natural numbers***. (And if you want to consider negative amounts as well, e.g. to talk about debts, you would use ***integers***.)
   - And you cannot [86] talk about the ***amount*** of dollar bills in your wallet.
   - If you have a units of a countable quantity such as dollar bills or coins, and b persons among whom you want to divide your a units equally, then the best you can do is give q units to each of the b persons, where q is the quotient of the division of a by b, and when you do that there will be a remainder of r undistributed dollar bills, where r is the remainder of the division of a by b.

[86] I really mean "you shouldn't, because it's wrong". Strictly speaking, you can say anything you want, in this free counrty of ours. But there are rules of grammar, and according to those rules it is wrong to say things like "a large amount of people were at the rally", or "she has a large amount of dollar bills". But it's O.K. to ta;lk about "a large amount of money". "People", like "dollar bills", or "coins", is countable. "Water", like "money", is uncountable.

- What happens if b is negative? Well, in this case you certainly cannot have $0 \leq r < b$, because if $b < 0$ this is impossible. But you can ask for a remainder r such that $0 \leq r < |b|$, where $|b|$ is the ***absolute value*** of b, that is, the number defined by

$$|x| = \begin{cases} x & \text{if} \quad x \geq 0 \\ -x & \text{if} \quad x < 0 \end{cases} . \tag{15.326}$$

- So the final condition is

$$a = bq + r \qquad \text{and} \quad 0 \leq r < |b| . \tag{15.327}$$

The division theorem says precisely that given integers a, b, there exist integers q, r such that (15.327) holds, provided, of course, that b is not equal to zero. And in addition it makes the very important and very useful assertion that q and r are ***unique***, that is, there is only one possible choice of q and r.

15.1.1 An example: even and odd integers

Example 70. Let us apply the division theorem to the case when $b = 2$. Suppose a is an integer.

What does the division theorem tell us about a?

The theorem makes two assertions, namely,

1. that the quotient and remainder exist (that's the ***existence part***),
2. that the quotient and remainder are unique (that's the ***uniqueness part***).

So let us look at each of these two parts, and see what it tells us about a.

The existence part of the theorem tells us that we can find integers q and r such that

$$a = 2q + r \text{ and } 0 \leq r < 2 .$$

Since $0 \leq r < 2$ and r is an integer, it follows that $r = 0$ or $r = 1$.

If $r = 0$ then $a = 2q$, so a is divisible by 2, that is, a is even.

If $r = 1$ then $a = 2q + 1$, so $a - 1 = 2q$, and then $a - 1$ is divisible by 2, that is, $a - 1$ is even, and, according to our definition of "odd", this implies that a is odd.

So we have shown that: either $r = 0$, in which case a is even, or $r = 1$, in which case a is odd. So ***the existence part of the division theorem tells us that a must be even or odd.***

The uniqueness part of the theorem tells us that we cannot find integers q, r such that

$$a = 2q + r \text{ and } 0 \leq r < 2\,,$$

and also find different integers q', r' such that

$$a = 2q' + r' \text{ and } 0 \leq r' < 2\,.$$

In particular, it is not possible to find integers q, q' such that

$$a = 2q \text{ and } a = 2q' + 1\,.$$

In other words, a cannot be both even and odd. So ***the uniqueness part of the division theorem tells us that*** a ***cannot be both even and odd.***

Summarizing: ***the division theorem, for*** $b = 2$***, tells us that an integer*** a ***has to be even or odd and cannot be both even and odd.*** And this is exactly Theorem 26, that we had to work so hard to prove!

In other words: ***The division theorem (that is, Theorem 56 below) is a generalization of the theorem that says that every integer is even or odd and not both.*** □

Now that we understand what the division theorem says for $b = 2$, let us look at what it says for general values of b.

- The division theorem says that, when you try to divide an integer a by 2, then one and only one of two things will happen:

  1. you will be able to divide a by 2 exactly, with a remainder equal to zero, and conclude that a is even,
  2. you will not be able to divide a by 2 exactly, but you will be able to do it with a remainder equal to 1, and conclude that $a - 1$ is divisible by 2, so a is odd.

  In other words, the division theorem, applied with $b = 2$, says exactly that that every integer is even or odd and not both. Furthermore:

  1. the *existence* part of the theorem says that every integer is even or odd;
  2. the *uniquenss* part of the theorem says that an integer cannot be both even and odd.

- The division theorem, applied with $b = 3$, says that, when you try to divide an integer a by 3, then one and only one of three things will happen:
  1. you will be able to divide a by 3 exactly, with a remainder equal to zero, and conclude that a is divisible by 3,
  2. you will not be able to divide a by 3 exactly, but you will be able to do it with a remainder equal to 1, and conclude that $a = 3q+1$ for some integer q, so $a - 1$ is divisible by 3.
  3. you will not be able to divide a by 3 exactly, but you will be able to do it with a remainder equal to 2, and conclude that $a = 3q+2$ for some integer q, so $a - 2$ is divisible by 3.
- The division theorem, applied with $b = 4$, says that, when you try to divide an integer a by 4, then one and only one of four things will happen: $4|a$, $4|a-1$, $4|a-2$, $4|a-3$.
- The division theorem, applied with $b = 5$, says that, when you try to divide an integer a by 5, then one and only one of five things will happen: $5|a$, $5|a-1$, $5|a-2$, $5|a-3$, $5|a-4$.
- $\cdots$
- The division theorem, applied with $b = 29$, says that, when you try to divide an integer a by 29, then one and only one of 29 things will happen: $29|a-j$ for $j \in \mathbb{Z}$, $0 \leq j < 29$.
- $\cdots$
- The division theorem, applied with $b = 372,508$, says that, when you try to divide an integer a by $372,508$, then one and only one of $372,508$ things will happen: $372,508|a-j$ for $j \in \mathbb{Z}$, $0 \leq j < 372,508$.

15.2 Precise statement of the division theorem

And here is, finally, the division theorem:

The division theorem for integers

Theorem 56 *If a, b are integers, and $b \neq 0$, then there exist unique integers q, r such that*

$$a = bq + r \text{ and } 0 \leq r < |b| .$$

15.2.1 The quotient and the remainder

Definition 24. *If a, b are integers, and $b \neq 0$, then the unique integers q, r such that*

$$a = bq + r \qquad \text{and} \qquad 0 \leq r < |b|$$

called, respectively, the quotient and the remainder of the division of a by b.

We use $\mathrm{QUO}(a, b)$ *and* $\mathrm{REM}(a, b)$ *to denote the quotient and the remainder of the division of a by b.* □

It follows from Definition 24 that, if $a \in \mathbb{Z}$, $b \in \mathbb{Z}$, and $b \neq 0$, then

1. $a = b \times \mathrm{QUO}(a, b) + \mathrm{REM}(a, b)$,
2. $\mathrm{QUO}(a, b) \in \mathbb{Z}$,
3. $\mathrm{REM}(a, b) \in \mathbb{Z}$ and $0 \leq \mathrm{REM}(a, b) < |b|$,
4. if q, r are integers such that $a = bq + r$ and $0 \leq r < |b|$, then $q = \mathrm{QUO}(a, b)$ and $r = \mathrm{REM}(a, b)$.

15.2.2 Some problems

Problem 83. ***Prove*** the following theorem.

Theorem 57. *If n is an integer, then there exist unique integers q, r such that*

$$n^2 = 4q + r \qquad \text{and} \qquad r = 0 \vee r = 1\,.$$

(HINT: First write $n = 4k + s$, with $0 \leq s < 4$, and then prove that $\text{REM}(n^2, 4)$ must be 0 or 1.) □

Problem 84. ***Prove*** the following theorem.

Theorem 58. *If m, n are integers, then there exist unique integers q, r such that*

$$m^2 + n^2 = 4q + r \qquad \text{and} \qquad r = 0 \vee r = 1 \vee r = 2\,.$$

(HINT: Use Theorem 57.) □

Problem 85. ***Prove*** that if $n = 3,409,583$, then there do not exist integers p, q such that $p^2 + q^2 = n$. □

15.3 Proof of the division theorem

The proof of the division theorem will be split up into two parts.

1. We will first prove the existence part. That is, we will prove

   Theorem 56.I. *If a, b are integers and $b \neq 0$ then there exist integers q, r such that $a = bq + r$ and $0 \leq r < |b|$.*

2. Then, after we have proved the existence result —i.e.., Theorem 56.I— - we will prove the uniqueness result. That is, we will prove

   Theorem 56.II. *If a, b are integers and $b \neq 0$, and q, r, q', r' are integers such that*

   $$a = bq + r \qquad \text{and} \qquad 0 \leq r < |b|\,,$$

   and

   $$a = bq' + r' \qquad \text{and} \qquad 0 \leq r' < |b|\,,$$

   then $q = q'$ and $r = r'$.

15.3.1 Proof of the existence part of the division theorem, using induction going forward and backward

We want to prove that

$$(\forall a\in\mathbb{Z})(\forall b\in\mathbb{Z})\Big(b\neq 0 \Longrightarrow (\exists q\in\mathbb{Z})(\exists r\in\mathbb{Z})(a=bq+r\wedge 0\leq r<|b|)\Big)\,. \quad (15.328)$$

This is logically equivalent[87] to

$$(\forall b\in\mathbb{Z})(\forall a\in\mathbb{Z})\Big(b\neq 0 \Longrightarrow (\exists q\in\mathbb{Z})(\exists r\in\mathbb{Z})(a=bq+r\wedge 0\leq r<|b|)\Big)\,. \quad (15.329)$$

and to

$$(\forall b\in\mathbb{Z})\Big(b\neq 0 \Longrightarrow (\forall a\in\mathbb{Z})(\exists q\in\mathbb{Z})(\exists r\in\mathbb{Z})(a=bq+r\wedge 0\leq r<|b|)\Big)\,. \quad (15.330)$$

So we are going to prove (15.330).

Let b be an arbitrary integer. We want to prove

$$b\neq 0 \Longrightarrow (\forall a\in\mathbb{Z})(\exists q\in\mathbb{Z})(\exists r\in\mathbb{Z})(a=bq+r\wedge 0\leq r<|b|)\,. \quad (15.331)$$

Assume that $b\neq 0$. We want to prove

$$(\forall a\in\mathbb{Z})(\exists q\in\mathbb{Z})(\exists r\in\mathbb{Z})(a=bq+r\wedge 0\leq r<|b|)\,. \quad (15.332)$$

Proposition (15.332) is a universal sentence whose universal quantifier is "$(\forall a\in\mathbb{Z})$". So we have the option of using induction forward and backward, and we are going to do it that way.

We let $P(a)$ be the statement

$$(\exists q\in\mathbb{Z})(\exists r\in\mathbb{Z})(a=bq+r\wedge 0\leq r<|b|)\,. \quad (15.333)$$

We are going to prove $(\forall a\in\mathbb{Z})P(a)$ by induction going forward and backward.

Basis step. We have to prove $P(0)$.

Clearly, $P(0)$ says that

$$(\exists q\in\mathbb{Z})(\exists r\in\mathbb{Z})(0=bq+r\wedge 0\leq r<|b|)\,, \quad (15.334)$$

which is an existential sentence.

[87] Two propositions A, B are logically equivalent if the proposition $A\iff B$ can be proved by pure logic, that is, using only the rules of logic. It is an easy exercise to see that, if $P(x,y)$ is a 2-variables predicate (that is, a sentence with the open variables x,y) and S is a set, then $(\forall x\in S)(\forall y\in S)P(x,y)$ is logically equivalent to $(\forall y\in S)(\forall x\in S)P(x,y)$.

To prove 15.334, we can use Rule $\exists_{prove}$, and for that purpose we have to produce witnesses, i.e., integers q, r such that

$$0 = bq + r \text{ and } 0 \leq r < |b| \,. \tag{15.335}$$

And this is very easy: just take $q = 0$, $r = 0$.

Then, with this choice of q, r, it is clear that (15.335) holds. (Notice that here we are using the fact that $b \neq 0$, to conclude that $r < |b|$.)

Hence (15.334) is true, and we have proved $\boxed{P(0)}$, and completed the basis step.

Inductive step. We have to prove that

$$(\forall a \in \mathbb{Z})\big(P(a) \iff P(a+1)\big) \,. \tag{15.336}$$

Let a be an arbitrary integer.

We want to prove "$P(a) \iff P(a+1)$".

For this purpose, we are going use Rule $\iff_{prove}$, and this requires that we prove the pair of implications "$P(a) \implies P(a+1)$" and "$P(a+1) \iff P(a)$".

Proof of "$P(a) \implies P(a+1)$".

Assume $P(a)$.

Then

$$(\exists q \in \mathbb{Z})(\exists r \in \mathbb{Z})(a = bq + r \wedge 0 \leq r < |b|) \,. \tag{15.337}$$

So we may pick integers q, r such that

$$a = bq + r \wedge 0 \leq r < |b| \,. \tag{15.338}$$

We want to prove that $P(a+1)$ is true, i.e., that

$$(\exists q \in \mathbb{Z})(\exists r \in \mathbb{Z})(a + 1 = bq + r \wedge 0 \leq r < |b|) \,, \tag{15.339}$$

and for that purpose we need witnesses q', r' for (15.339).

The most obvious choice would be to take $q' = q$, $r' = r + 1$.

Then $a + 1 = bq' + r'$. But we cannot prove that $r' < |b|$, because from "$r < |b|$" all we can conclude is that $r' \leq |b|$, and that's not what we need.

On the other hand, if we knew that $r + 1 < |b|$, then the conclusion "$r' < |b|$" would follow.

Therefore, so we have found the desired witnesses q', r', and proved (15.339), under the assumption that $r + 1 < |b|$.

We now have to take care of the possibility that $r + 1 \geq |b|$.

In that case, since $r < |b|$, it follows that $r + 1 = |b|$. (REASON: r and $|b|$ are integers, so if $r < |b|$ then $r + 1 \leq |b|$. So if $r + 1 \geq |b|$ it follows that $r + 1 = |b|$.)
Then

$$\begin{aligned} a + 1 &= bq + r + 1 \\ &= bq + |b| \,. \end{aligned}$$

Define an integer μ as follows:

$$\mu = \begin{cases} 1 & \text{if} \quad b > 0 \\ -1 & \text{if} \quad b < 0 \,. \end{cases}$$

Then $|b| = \mu b$.
Therefore

$$\begin{aligned} a + 1 &= bq + |b| \\ &= bq + \mu b \\ &= b(q + \mu) \\ &= b(q + \mu) + 0 \,. \end{aligned}$$

Hence, if we choose our witnesses q', r' by letting $q' = q + \mu$ and $r' = 0$, it follows that $a + 1 = bq' + r'$ and $0 \leq r' < |b|$. So we have also proved (15.339).
So we have proved (15.339) in both cases, when $r + 1 < |b|$ and when $r + 1 \geq |b|$.
Hence we have proved (15.339), i.e., we have proved $P(a + 1)$.
Since we have proved $P(a + 1)$ assuming $P(a)$, it follows from Rule $\Longrightarrow_{prove}$ that we have proved $\boxed{P(a) \Longrightarrow P(a + 1)}$.
Proof of "$P(a + 1) \Longrightarrow P(a)$".
Assume $P(a + 1)$.
Then

$$(\exists q \in \mathbb{Z})(\exists r \in \mathbb{Z})(a + 1 = bq + r \wedge 0 \leq r < |b|) \,, \qquad (15.340)$$

So we may pick integers q, r such that

$$a + 1 = bq + r \wedge 0 \leq r < |b| \,. \qquad (15.341)$$

We want to prove that $P(a)$ holds, i.e., that

$$(\exists q \in \mathbb{Z})(\exists r \in \mathbb{Z})(a = bq + r \wedge 0 \leq r < |b|) \,, \qquad (15.342)$$

and for that purpose we need witnesses q', r' for (15.342).
The most obvious choice is to take $q' = q$, $r' = r - 1$.

Then $a = bq'+r'$ (because $a+1 = bq+r$, $q' = q$, and $r' = r-1$). And $r' < |b|$ (because $r < |b|$ and $r' < r$.) But we cannot prove that $r' \geq 0$, because r could be 0, in which case r' would be -1.

On the other hand, if we knew that $r > 0$, then the conclusion "$0 \leq r' < |b|$" follows.

So we have found the desired witnesses q', r', and proved (15.342), under the assumption that $r > 0$.

We now have to take care of the possibility that $\boxed{r = 0}$.

In that case, we have

$$\begin{aligned} a+1 &= bq + r \\ &= bq\,. \end{aligned}$$

So

$$\begin{aligned} a &= bq - 1 \\ &= bq - |b| + |b| - 1 \\ &= bq - \mu b + |b| - 1 \\ &= b(q-\mu) + |b| - 1\,, \end{aligned}$$

where μ is the number defined earlier.

Hence, if we choose our witnesses q', r' by letting $q' = q - \mu$ and $r' = |b| - 1$, it follows that $a = bq' + r'$ and $0 \leq r' < |b|$. So we have also proved (15.342).

So we have proved (15.342) in both cases, when $r > 0$ and when $r = 0$.

Hence we have proved (15.342), i.e., we have proved $P(a)$.

Since we have proved $P(a)$ assuming $P(a+1)$, it follows from Rule $\Longrightarrow_{prove}$ that we have proved $\boxed{P(a+1) \Longrightarrow P(a)}$.

Since we have proved the implications $P(a) \Longrightarrow P(a+1)$ and $P(a+1) \Longrightarrow P(a)$, it follows from Rule $\Longleftrightarrow_{prove}$ that we have proved $\boxed{P(a) \Longleftrightarrow P(a+1)}$.

Since we have proved $P(a+1) \Longleftrightarrow P(a)$ for an arbitrary integer a, it follows from Rule $\forall_{prove}$ that

$$(\forall a \in \mathbb{Z})\big(P(a) \Longleftrightarrow P(a+1)\big)\,. \tag{15.343}$$

Then the principle of mathematical induction going forward and backward implies that $(\forall a \in \mathbb{Z})P(a)$, that is,

$$(\forall a \in \mathbb{Z})(\exists q \in \mathbb{Z})(\exists r \in \mathbb{Z})(a = bq + r \wedge 0 \leq r < |b|)\,. \tag{15.344}$$

Proposition (15.344) was proved under the assumption that $b \neq 0$. Hence

$$b \neq 0 \Longrightarrow \Big((\forall a \in \mathbb{Z})(\exists q \in \mathbb{Z})(\exists r \in \mathbb{Z})(a = bq + r \wedge 0 \leq r < |b|) \Big) . \quad (15.345)$$

Proposition (15.345) was proved for an arbitrary integer b. Hence

$$(\forall b \in \mathbb{Z}) \Bigg(b \neq 0 \Longrightarrow \Big((\forall a \in \mathbb{Z})(\exists q \in \mathbb{Z})(\exists r \in \mathbb{Z})(a = bq + r \wedge 0 \leq r < |b|) \Big) \Bigg) . \quad (15.346)$$

And this completes the existence proof. **Q.E.D.**

15.3.2 Proof of the uniqueness part of the division theorem

We now prove, finally,

Theorem 56.II. *If a and b are integers and $b \neq 0$, then the integers q, r such that $a = bq + r$ and $0 \leq r < |b|$ are unique.*
Proof.

Let $a \in \mathbb{Z}$ be arbitrary. Let $b \in \mathbb{Z}$ be arbitrary.

Assume that $b \neq 0$. We want to prove that

(*) If q, q', r, r' are integers such that

$$\begin{aligned} a &= bq + r , && (15.347)\\ 0 &\leq r < b , && (15.348)\\ a &= bq' + r' , && (15.349)\\ 0 &\leq r' < b , && (15.350) \end{aligned}$$

then $q = q'$ and $r = r'$.

Let q, q', r, r' be integers such that (15.347), (15.348), (15.349), and (15.350) hold.

We will prove that $q = q'$ and $r = r'$.

Without loss of generality, we may assume that $r \geq r'$. (Reason: if r was $< r'$, just change the names of r, r' and call them r' and r.)

Then

$$0 \leq r - r' < b . \quad (15.351)$$

(Reason: 0 is $\leq r - r'$ because $r \geq r'$. And $r - r' < b$ because $r - r' \leq r$, since $r' \geq 0$, and $r < b$.)

On the other hand, $a = bq + r$ and $a = bq' + r'$, so

$$bq + r = bq' + r' .$$

Therefore

$$b(q' - q) = r - r' \,. \tag{15.352}$$

Then

$$|b| \cdot |q' - q| = |r - r'| \,, \tag{15.353}$$

because $|xy| = |x| \cdot |y|$ for arbitrary real numbers x, y.

Since q and q' are integers, the number $|q - q'|$ is a nonnegative integer.

We now prove[88] that $q = q'$.

Assume that $q \neq q'$.

Then the nonnegative integer $|q - q'|$ is not zero, so it is a natural number.

And then $|q - q'| \geq 1$.

Therefore (15.353) implies that $|r - r'| \geq |b|$.

But $r - r' \geq 0$, because $r \geq r'$.

Hence $|r - r'| = r - r'$.

It follows that $r - r' \geq |b|$.

So it's not true that $r - r' < |b|$.

But (15.351) tells us that $r - r' < |b|$.

So $r - r' < |b|$ and $\sim r - r' < |b|$, which is a contradiction.

This proves that $\boxed{q = q'}$.

And then (15.353) implies that $\boxed{r = r'}$.

So we have proved (**), for arbitrary integers a, b such that $b \neq 0$.

This completes the proof of the uniqueness part. So our proof is complete.**Q**.**E**.**D**.

[88] by contradiction , naturally.

16 The Well-ordering Principle

The ***well-ordering principle*** (WOP) is a very simple consequence of the PMI, and is a very powerful tool for proving properties of the integers.

16.1 Statement of the Well-ordering Principle

In order to state the WOP, we need to know what is meant by "a smallest member" of a set of integers (or of real numbers).

16.1.1 The smallest member of a set of integers

Definition 25. *If S is a subset*[89] *of*[90] *$\mathbb{Z}$, a smallest member of S is a member s of S such that*

$$(\forall t \in S) s \leq t\,. \tag{16.354}$$

Example 71.

- Let S be the set of all prime numbers. Then S is a set of natural numbers, and the smallest member of S is 2.
- Let S be the set of all integers. Then S has no smallest member. (Proof: suppose s was a smallest member of S. Then $s \in \mathbb{Z}$, and $s \leq t$ for every integer t. In particular, $s \leq t-1$. But $s > t-1$, so $s \leq t-1 \wedge \sim s \leq t-1$, which is a contradiction.)
- Let S be the set of all real numbers. Then S has no smallest member. (Proof: suppose s was a smallest member of S. Then $s \in \mathbb{R}$, and $s \leq t$ for every real number t. In particular, $s \leq t-1$. But $s > t-1$, so $s \leq t-1 \wedge \sim s \leq t-1$, which is a contradiction.)
- Let S be the set of all nonnegative real numbers. In other words, $S = \{x \in \mathbb{R} : x \geq 0\}$. Then the smallest member of S is 0.
- Let S be the set of all positive real numbers. In other words, $S = \{x \in \mathbb{R} : x > 0\}$.) Then S has no smallest member. (Proof: suppose s was a smallest member of S. Then $s \in \mathbb{R}$, $s > 0$, and $s \leq t$ for every $t \in \mathbb{R}$ such that $t > 0$. In particular, $s \leq \frac{s}{2}$, so $2s \leq s$. But $2s > s$, because $2 > 1$ and we can multiply both sides by s, since $s > 0$, thus getting $2s > s$. So $2s \leq s \wedge \sim 2s \leq s$, which is a contradiction.)□

[89] The meaning of "subset" is discussed in section 4.1.7, on page 44.
[90] or of $\mathbb{Q}$, or of $\mathbb{R}$, or of any set equipped with an order relation $\leq$.

16.1.2 Uniqueness of the smallest member of a set

In Definition 25 we explained what it means for an integer s to be ***a*** smallest member of a set S of integers.

Can we talk about ***the*** smallest member of S?

The answer to this type of question, in general, is that

- we can talk about ***the*** ZZZ if there exists only one ZZZ,
- we cannot talk about ***the*** ZZZ, and we talk instead about ***a*** ZZZ, if there exist more than one ZZZ.

Example 72.

1. We can say that Jeff Bezos[91] is ***the*** richest person in the world, because there is only one richest person in the world.
2. But we do not say that Michael Bloomberg[92] is ***the*** billionaire, because there are lots of billionaires; we can say that Michael Bloomberg is ***a*** billionaire.
3. We can say that 2 is ***the*** smallest prime number, because there is only one smallest prime number.
4. But we do not say that 2 is ***the*** prime number, because there are lots of prime numbers; we say that 2 is ***a*** prime number. □

The following theorem states the obvious fact that if a set has a smallest member, then that smallest member is unique. This is a completely obvious fact, but in Mathematics everything has to be proved. If is is obvious, then there should exist a simple proof.

Trivial theorem. *If a subset S of $\mathbb{Z}$ (or of $\mathbb{Q}$, or of $\mathbb{R}$) has a smallest member, then it has only one smallest member.*

Proof. Let s_1, s_2 be smallest members of S. Since s_1 is a smallest member of S, $s_1 \leq t$ for every $t \in S$. In particular, since $s_2 \in S$, $s_1 \leq s_2$.

Similarly, $s_2 \leq s_1$. So $s_1 = s_2$. **Q.E.D.**

[91] Or maybe it's Bill Gates?

[92] Or Donald Trump, whoever you like best

16.1.3 Statement of the Well-ordering Principle: The Standard Version

The standard version of the well-ordering principle, the one that you will find in most textbooks, says that *every nonempty set of natural numbers has a smallest member:*

THE WELL-ORDERING PRINCIPLE (WOP) STANDARD VERSION

Theorem 59 *Every nonempty set of natural numbers has a smallest member.*

In formal language, Theorem 59 says that

$$(\forall S)\Big(\big(S \subseteq \mathbb{N} \wedge S \neq \emptyset\big) \Longrightarrow (\exists s \in S)(\forall t \in S) s \leq t\Big), \quad (16.355)$$

16.1.4 Sets that are bounded below

Definition 26. *A subset S of $\mathbb{Z}$ is <u>bounded below</u> if there exists an integer s_* such that $S \subseteq \mathbb{Z}_{\geq s_*}$, i.e., that $(\forall s \in S) s \geq s_*$.* □

Definition 27. *A subset S of $\mathbb{Z}$ is <u>bounded above</u> if there exists an integer s_* such that $(\forall s \in S) s \leq s_*$.* □

So a set S of integers is bounded below if there is an integer s_* such that all the members of S are to the right[93] of s_*.

And, similarly, a set S of integers is bounded above if there is an integer s_* such that all the members of S are to the left of s_*. "

[93] Let us be precise: "to the right of" means "$\geq$"; "to the left of" means "$\leq$"; "strictly to the right of" means "$>$"; and "strictly to the left of" means "$<$".

Problem 86. ***Prove*** that a set S of integers is bounded above if and only if the set $-S$ given by

$$-S = \{s \in \mathbb{Z} : -s \in S\}$$

is bounded below.

16.1.5 Statement of the Well-ordering Principle: A More General Version

There is a sligthly more general version that is often more useful than the standard one: instead of subsets of $\mathbb{N}$, we can consider equally well sets that are subsets of $\mathbb{Z}_{\geq s_*}$ for some $s_* \in \mathbb{Z}$. (Recall that, if $s_* \in \mathbb{Z}$, then $\mathbb{Z}_{\geq s_*}$ is the set of all integers n such that $n \geq s_*$. That is,

$$\mathbb{Z}_{\geq s_*} = \{n \in \mathbb{Z} : n \geq s_*\}. \tag{16.356}$$

as explained earlier in these notes.)

We are now ready to state the WOP:

THE WELL-ORDERING PRINCIPLE (WOP) GENERAL VERSION

Theorem 60 *Every nonempty set of integers which is bounded below has a smallest member.*

In formal language, Theorem 60 says that

$$(\forall s_* \in \mathbb{Z})(\forall S)\Big(\big(S \subseteq \mathbb{Z}_{\geq s_*} \wedge S \neq \emptyset\big) \Longrightarrow (\exists s \in S)(\forall t \in S) s \leq t\Big). \tag{16.357}$$

16.2 Proof of the Well-Ordering Principle

We want to prove (16.357). So we fix an arbitrary integer s_*, and try to prove that

$$(\forall S)\bigg((S \subseteq \mathbb{Z}_{\geq s_*} \wedge S \neq \emptyset) \Longrightarrow (\exists s \in S)(\forall t \in S) s \leq t\Big)\,. \tag{16.358}$$

We will first prove a lemma.

Lemma. *If $n \in \mathbb{Z}$, $S \subseteq \mathbb{Z}_{\geq s_*}$ and $n \in S$, then S ha a smallest member.*

In formal language, the lemma says:

$$(\forall n \in \mathbb{Z}_{\geq s_*})(\forall S)\bigg(\Big(S \subseteq \mathbb{Z}_{\geq s_*} \wedge n \in S\Big) \Longrightarrow (\exists s \in S)(\forall t \in S) s \leq t\bigg)\,. \tag{16.359}$$

Before we prove the lemma, let us show how Theorem 60 —i.e., formula (16.358— follows immediately from it.

Proof of formula (16.358) using the lemma:
Let S be a nonempty subset of $\mathbb{Z}_{\geq s_*}$. Since $S \neq \emptyset$, we may pick a member n of S. Then $S \subseteq \mathbb{Z}_{\geq s_*}$ and $n \in S$. So by the lemma, S has a smallest member, This proves Theorem 60.

Proof of the lemma.
We will do a proof by induction starting at s_*.

In the proof, we will write $H(S)$ for "S has a smallest member".
Let $P(n)$ be the predicate "for every subset S of $\mathbb{Z}_{\geq s_*}$ such that $n \in S$ has a smallest member". That is, $P(n)$ stands for

$$(\forall S)\bigg(\Big(S \subseteq \mathbb{Z}_{\geq s_*} \wedge n \in S\Big) \Longrightarrow H(S)\bigg)\,. \tag{16.360}$$

We will prove $(\forall n \in \mathbb{Z}_{\geq s_*})P(n)$, which is exactly formula (16.359), by induction starting with $n = s_*$.

Basis step. We have to prove $P(s_*)$. But $P(s_*)$ says "if $S \subseteq \mathbb{Z}_{\geq s*}$ and $s_* \in S$, then $H(S)$". And this is obvious because if $s_* \in S$ and $S \subseteq \mathbb{Z}_{\geq s_*}$, then all the members of S are $\geq s_*$, so s_* is the smallest member of S, and then $H(S)$ is true. Hence $\boxed{P(s_*)}$ holds.

Inductive srep. We have to prove

$$(\forall n \in \mathbb{Z}_{\geq s*})\big(P(n) \Longrightarrow P(n+1)\big)\,. \tag{16.361}$$

Let $n \in \mathbb{Z}_{\geq s_*}$ be arbitrary.

We want to prove the implication $P(n) \Longrightarrow P(n+1)$.

Assume $P(n)$. We want to prove that $P(n+1)$.

But $P(n+1)$ says "if S is an arbitrary subset of $\mathbb{Z}_{\geq s*}$ such that $n+1 \in S$, then $H(S)$".

Let S be an arbitrary subset of $\mathbb{Z}_{\geq s_*}$ such that $n+1 \in S$. We want to prove that $H(S)$.

There are two possibilities, namely, $\boxed{n \in S \text{ or } n \notin S}$.

We first consider the case when $n \in S$.

Assume $\boxed{n \in S}$.

Then, since we are assuming that $P(n)$ holds, $\boxed{H(S)}$.

So $\boxed{n \in S \Longrightarrow H(S)}$. [Rule $\forall_{prove}$]

We next consider the case when $n \notin S$.

Assume $\boxed{n \notin S}$.

Let [94] $T = S \cup \{n\}$.

Then $T \subseteq \mathbb{Z}_{\geq s_*}$, because $S \subseteq \mathbb{Z}_{\geq s_*}$ and $n \in \mathbb{Z}_{\geq s_*}$.

Furthermore, $n \in T$. Since we are assuming that $P(n)$ holds, and $P(n)$ says "if a subset of $\mathbb{Z}_{\geq s_*}$ contains n, then the subset has a smallest member", it follows that $H(T)$.

Let $\bar{t}$ be the smallest member of T. Then

$$\bar{t} \in T \wedge (\forall t \in T)\bar{t} \leq t\,. \tag{16.362}$$

In particular, since $S \subseteq T$, (16.362) implies

$$(\forall t \in S)\bar{t} \leq t\,, \tag{16.363}$$

So $\bar{t}$ is less than or equal to every member of S.

If $\boxed{\bar{t} \in S}$, then $\bar{t}$ is the smallest member of S, so $\boxed{H(S)}$.

If $\boxed{\bar{t} \notin S}$, then $\bar{t} = n$, because $T = S \cup \{n\}$ and $\bar{t} \in T$.

Furthermore, every member t of S satisfies $t \geq \bar{t}$, by (16.362).

So, if $t \in S$ then $t \geq \bar{t}$, i.e., $t \geq n$, but t cannot be equal to n, because $t \in S$ and $n \notin S$ (since $n = \bar{t}$ and $\bar{t} \notin S$). Hence $t > n$, and then $t \geq n+1$.

So we have proved that every member t of S satisfies $t \geq n+1$.

Since $n+1 \in S$, it follows that $n+1$ is the smallest member of S, so $\boxed{H(S)}$.

Since we have shown that $H(S)$ both when $\bar{t} \in S$ and when $\bar{t} \notin S$, we have proved $\boxed{\boxed{H(S)}}$.

[94] That is, T is the set obtained from S by adding n as a new member to S.

Since we have proved $H(S)$ assuming $n \notin S$, we can conclude that $\boxed{n \notin S \Longrightarrow H(S)}$.

Since we have proved $\boxed{n \in S \Longrightarrow H(S)}$ and $\boxed{n \notin S \Longrightarrow H(S)}$, it follows that $\boxed{H(S)}$.

So we have proved $H(S)$ for an arbitrary subset S of $\mathbb{Z}_{\geq s_*}$ such that $n+1 \in S$. And this proves that $\boxed{P(n+1)}$ holds.

So we have proved $P(n+1)$ assuming $P(n)$. Hence $\boxed{P(n) \Longrightarrow P(n+1)}$.

And, since we proved that $P(n) \Longrightarrow P(n+1)$ for arbitrary $n \in \mathbb{Z}_{\geq s_*}$, it follows that

$$(\forall n \in \mathbb{Z}_{\geq s)*}\big(P(n) \Longrightarrow P(n+1)\big)\,. \tag{16.364}$$

which is exactly (16.361).

This completes the inductive step. Since we have also carried out the basic step, the PMI enables us to conclude that

$$(\forall n \in \mathbb{Z}_{\geq s)*} P(n)\,, \tag{16.365}$$

which is exactly the statement of the lemma.

So we have proved the lemma and, as explained above, Theorem 60 is proved.

Problem 87. A largest member of a set S of integers (or of real numbers) is a member s of S such that $(\forall t \in S) t \leq s$.

Prove the following theorems:

Theorem 61. *If a set S of integers (or of real numbers) has a largest member, then this largest member is unique.*

Theorem 62. *Every nonempty set of integers which is bounded above has a largest member.*

HINT for Theorem 62: use the well-ordering principle (Theorem 60) and the result of Problem 86. □

16.3 A simple example of a proof using well-ordering: existence of prime factors

As an illustration of the power of the well-ordering pincple, let us use it to prove the following

Theorem 63. *If n is any natural number such that $n > 1$, then n has a prime factor. (That is, there exists a prime number p such that p is a factor of n, i.e., equivalently, $p|n$.)*

Idea of the proof. Let $n \in \mathbb{N}$ be arbitrary. Assume that $n > 1$. Then n has at least one nontrivial[95] natural number factor m. (Reason: n itself is one such factor.)

Let p be smallest of all the nontrivial natural number factors of n. Then p must be prime, because if p was not prime then p would have a smaller nontrivial factor q, and then q would be a nonrivlal natural number factor of n smaller than p.

And now we write this down in a more detailed fashion.

Proof.
Let n be a natural number such that $n > 1$.
Let F be the set of all natural numbers m such that $m > 1$ and m is a factor of n.
Then F is nonempty. (Proof: The number n is obviously a factor of n. And $n > 1$. So $n \in F$.)
Also, F is a subset of $\mathbb{Z}$, and F is bounded below (because $F \subseteq \mathbb{N}$).
By the well-ordering principle, F has a smallest member.
Let q be the smallest member of F.
Then q is a factor of n, and $q > 1$.
Furthermore, we claim that q is prime.

> *Proof that q is prime.*
>
> Suppose q was not prime.
> Then either $q = 1$ or q has a natural number factor other than 1 and q.
> Pick one such factor and call it r.
> Then r is a factor of q, so $q = rk$ for some natural number k.
> And q is a factor of n, so $n = qj$ for some natural number j.
> So $n = qj = (rk)j = r(jk)$.
> So r is a factor of n.
> But $r < q$, because r is a factor of q and r is not q.
> And $r > 1$, because r is a factor of q and r is not 1.
> Since r is factor of n and $r > 1$, it follows that $r \in F$.
> Since $r < q$ and $r \in F$, q is not the smallest member of F.
> But q is the smallest member of F.
> So we have reached a contradiction.
> So q is prime.

Hence q is a prime number which is a factor of n. So n has a prime factor. Q.E.D.

[95] "Nontrivial" means "not equal to 1".

16.4 More examples of simple proofs using well-ordering

Every proof that can be done by induction can also be done using well ordering. Indeed, suppose $P(n)$ is a one-variable predicate, and you can prove $P(1)$ and $(\forall n \in \mathbb{N})\big(P(n) \implies P(n+1)\big)$.

Then, instead of invoking the PMI, you could argue by well-ordering as follows. Call a natural number "bad" if $P(n)$ is not true. We want to prove that there are no bad numbers. Let B be the set of all bad natural numbers. We want to prove that B is empty. Suppose B is not empty. Then by the WOP B has a smallest member b. So b is bad but every natural number c such that $c < b$ is good (i.e., not bad). Then b cannot be 1, because $P(1)$ is true, so 1 is good. Since $b \in \mathbb{N}$ and $b \neq 1$, $b-1$ is a natural number. And $b-1$ is not bad, because b is the smallest bad natural number. So $b-1$ is good, that is, $P(b-1)$ is true. But then, since the implication $P(n) \implies P(n+1)$ is true for every $n \in \mathbb{N}$, it is true for $n = b-1$, which means that $P(b-1) \implies P(b)$ is true. Since $P(b-1)$ is true, t follows that $P(b)$ is true. So b is good, and we have derived a contradiction. Hence $B = \emptyset$.

Example 73. Let us prove using well-ordering that *if n is natural number, then $8^n - 5^n$ is divisible by* 3.

(We have already proved this by induction. I want to show that it can be done using well-ordering, and it's almost the same proof.)
Proof. We want to prove that

$$(\forall n \in \mathbb{N}) 3 | 8^n - 5^n \,. \tag{16.366}$$

Call a natural number n "bad" if 3 does not divide $8^n - 5^n$.

Ler B be the set of all bad natural numbers. We want to prove that $B = \emptyset$.

> Assume that $B \neq \emptyset$.
> Then, by the WOP, B has a smallest member b.
> Then b is bad, so $8^b - 5^b$ is not divisible by 3.
> In paericular, this means that $b \neq 1$, because $8^1 - 5^1$ is divisible by 3.
> So $b-1$ is a natural number, and $8^{b-1} - 5^{b-1}$ is divisible by 3.
> So we can write
> $$8^{b-1} - 5^{b-1} = 3k \,, \;\; k \in \mathbb{Z} \,.. \tag{16.367}$$
>
> Then
> $$8 \times (8^{b-1} - 5^{b-1}) = 3 \times 8k \,. \tag{16.368}$$

So

$$8^b - 8 \times 5^{b-1} = 3 \times 8k\,, \tag{16.369}$$

and then

$$8^b = 8 \times 5^{b-1} + 3 \times 8k\,, \tag{16.370}$$

But $8 = 5 + 3$, so

$$8 \times 5^{b-1} = 5 \times 5^{b-1} + 3 \times 5^{b-1} = 5^b + 3 \times 5^{b-1}\,, \tag{16.371}$$

so

$$8^b = 5^b + 3 \times 5^{b-1} + 3 \times 8k\,, \tag{16.372}$$

and then

$$8^b = 5^b + 3(5^{b-1} + 8k)\,, \tag{16.373}$$

so that

$$8^b - 5^b = 3(5^{b-1} + 8k)\,, \tag{16.374}$$

Let $j = 5^{b-1} + 8k$. Then $j \in \mathbb{Z}$ and

$$8^b - 5^b = 3j\,. \tag{16.375}$$

Hence $3|8^b - 5^b$. That is, $\boxed{b \text{ is good}}$.
But b is bad. So we have arrived at a contradiction.

The contradiction arose from assuming that B was nonemoty.

Hence B is empty, and our theorem is proved. **Q**.**E**.**D**.

Problem 88. ***Prove***, using well-ordering, that

$$(\forall n \in \mathbb{N})5|8^n - 3^n\,. \tag{16.376}$$

HINT: Use the same method as in the proof of statement (16.366), in Eaxmple 73. □

16.5 An example of a proof using well-ordering: proof of the existence of a coprime representation (a.k.a. "coprime expression", or "irreducible representation") of a rational number

If r is a rational numnber, the definition of "rational number" tells us that there exist integers m, n such that $n \neq 0$ and $r = \frac{m}{n}$.

It turns out that we can always pick m, n in such a way that the following facts are also true:

CR1 $n > 0$,

CR2 $m \perp n$.

and, of course,

CR3 $r = \frac{m}{n}$.

Remark 19. What does "$m \perp n$" mean? It means "m and n are coprime".

The definition of "coprime integers" has been given before, on page 64. (See Definition 15.) □

Definition 28. *Let r be a rational number. A coprime representation of r is a pair m, n of integrers such that conditions CR1, CR2, CR3 hold.*

Example 74.

- If $r = \frac{48}{18}$, then a coprime representation of r is given by writing $r = \frac{8}{3}$.
- If $r = \frac{-3}{-2}$, then a coprime representation of r is given by writing $r = \frac{3}{2}$.
- If $r = \frac{3}{-2}$, then a coprime representation of r is given by writing $r = \frac{-3}{2}$. □

The precise statement of the result announced above is as follows:

Theorem 64. *If r is a rational number, then there exists a coprime representation of r. That is,*

$$(\forall r \in \mathbb{Q})(\exists m \in \mathbb{Z})(\exists n \in \mathbb{Z})\left(n > 0 \wedge m \perp n \wedge r = \frac{m}{n}\right). \tag{16.377}$$

Proof.

Let r be an arbitrary rational number.

Let

$$S = \left\{ n \in \mathbb{N} : rn \in \mathbb{Z} \right\}. \tag{16.378}$$

Then, clearly, $S \subseteq \mathbb{N}$. (That is, S is a set of natural numbers.)

Claim: S is not empty.

Proof: Since r is rational, we can write $r = \frac{a}{b}$, where a and b are integers and $b \neq 0$.

Let

$$\begin{aligned} n &= \begin{cases} b & \text{if} \quad b > 0\,, \\ -b & \text{if} \quad b < 0\,, \end{cases} \\ m &= \begin{cases} a & \text{if} \quad b > 0\,, \\ -a & \text{if} \quad b < 0\,. \end{cases} \end{aligned}$$

Then $m \in \mathbb{Z}$, $n \in \mathbb{Z}$, $n > 0$, and $r = \frac{m}{n}$.

Hence $rn = m$, so $rn \in \mathbb{Z}$.

Furthermore, $n \in \mathbb{N}$.

Hence $n \in S$.

So S is not empty. *This completes the proof of the claim.*

Since S is a nonempty set of natural numbers, the Well-ordering principle tells us that S has a smallest member. (Furthermore, we know from the "Trivial Theorem" on page 308 that the smallest member of a nonempty set of real numbers, when it exists, is unique, so we can talk about ***the*** smallest member of S.)

Let q be the smallest member of S.

Let $p = rq$.

Then p is an integer, because $q \in S$.

Also, $q \in \mathbb{N}$ (because $q \in S$).

And $r = \frac{p}{q}$, because $rq = p$.

Claim: $p \perp q$.

Proof: Suppose $\sim p \perp q$.

Then p and q have a common factor $k \in \mathbb{Z}$ such that $k > 1$.

Siince $k|p$ and $k|q$, we may write

$$p = ku \text{ and } q = kv\,, \text{ where } u \in \mathbb{Z} \wedge v \in \mathbb{Z}\,.$$

Then $v \in \mathbb{Z}$ and $v > 0$ (because $kv = q$, $k > 0$, and $q > 0$).

Also, $r = \frac{p}{q} = \frac{ku}{kv} = \frac{u}{v}$.

Therefore $rv = u$.

Since $rv = u$, $u \in \mathbb{Z}$, and $v \in \mathbb{N}$, it follows that $v \in S$.

But $v < q$, because $q = kv$ and $k > 1$.

Since $v \in S$ and $v < q$, it follows that q is not the smallest member of S.

So *q is the smallest member of S and q is not the smallest member of S.*

So we have derived a contradiction from the assumption that p and q are not coprime.

Hence p and q are coprime.

So $r = \frac{p}{q}$, $p \in \mathbb{Z}$, $q \in \mathbb{Z}$, $q > 0$, and $p \perp q$.

Therefore, we have found a coprime representation of r.

So r has a coprime representation.

We have proved that *if r is an arbitary rational number then r has a coprime representaion.* This completes our proof. **Q.E.D.**

16.5.1 Is the coprime representation unique?

We have just proved that every rational number r has a coprime representation.

Is that coprime representation unique?

It turns out that this is an important question. We will see in Section 18.5 that the uniqueness of the coprime representation has important applications. So we should be able to prove that the coprime represenation of a rational number is unique.

Unfortunately, we cannot do it right now. We need a new technique, called ***Bézout's lemma***, which will enable us to prove ***Euclid's lemma***. These results will be proved in Sections 17 and 18.

And then, armed with these powerfull tools, we will be able to prove the uniqueness of the coprime representation in section 18.4 and then use it in section 18.5.

16.6 Another example of a proof using well-ordering: a second proof of the existence part of the fundamental theorem of arithmetic

In this section we prove the existence part of the ***fundamental theorem of arithmetic (FTA).*** This theorem is one of the most important results in integer arithmetic. It says that every natural number n such that $n \geq 2$ can be written as a product of primes in a unique way. (That is, not only is the number equal to a product of primes, but there is only one way to write it as a product of primes.) We will prove a part of the FTA, namely, the assertion that if $n \in \mathbb{N}$ and $n \geq 2$ then n can be written as a product of primes.

The proof of uniqueness requires more sophisticated tools, and will be done later.

Theorem 65. *Every natural number n such that $n \geq 2$ is a product of prime numbers*

Remark 20. The precise meaning of "is a product of prime numbers" was dicussed in detail earlier, in section 2.3.3, on page 22.

In particular, as we explained there, ***a single prime number is a product of prime numbers***. So, for example, 2, 3, 5, 7, ***are*** products of prime numbers. □

16.6.1 Outline of the strategy for proving the theorem

Call a natural number n "bad" if $n > 1$ and n is not a product of primes.

What we want is to prove is that there are no bad natural numbers.

The strategy is going to be this: we let B be the set of all bad numbers, so our goal is to prove that B is empty. For this purpose, we assume it is nonempty, and use the well-ordering Principle to conclude that it has a smallest member b. Then b is bad, and in addition b is the smallest bad natural number. But then b cannot be prime, because if it is prime then it is a product of primes, so b would not be bad. Since $b > 1$, and b is not

prime, b must be a product cd of two smaller natural numbers. But then c and d cannot be bad. So c is a product $p_1 \times p_2 \times \cdots \times p_k$ of primes, and d is a product $q_1 \times q_2 \times \cdots \times q_j$ of primes. So

$$b = cd = p_1 \times p_2 \times \cdots \times p_k \times q_1 \times q_2 \times \cdots \times q_j \,.$$

But then b is a product of primes, so b is not bad. But b is bad, and we got a contradiction. Hence B is empty, and that means that there are no bad numbers.

16.6.2 The proof

Let B be the set of all natural numbers n such that $n \geq 2$ and n is not a product of primes.

We want to prove that the set B is empty. For this purpose, we assume that B is not empty and try to get a contradiction.

So assume that $B \neq \emptyset$. By the well-ordering principle, B has a smallest member b. Then $b \in B$, so

a. b is a natural number,

b. $b \geq 2$,

c. $\boxed{b \text{ is not a product of primes}}$.

And, in addition,

d. b is the smallest member of B, that is,

$$(\forall m)(m \in B \Longrightarrow m \geq b) \,.$$

Since b is not a product of primes, it follows in particular that b is not prime. (Reason: if b was prime, then b would be a product of primes according to our definition.)

Since b is not prime, there are two possibilities: either $b = 1$ or b has a factor k which is a natural number such that $k \neq 1$ and $k \neq b$.

But the fist possibility ($b = 1$) cannot arise, because $b \geq 2$.

Hence the second possibility occurs. That is, we can pick a natural number k such that k divides b, $k \neq 1$, and $k \neq b$.

Since $k|b$, we can pick an integer j such that

$$b = jk\,.$$

And then j has to be a natural number. (Reason: we know that $k \in \mathbb{N}$, so $k > 0$. If j was ≤ 0, it would follow that $kj \leq 0$. But $kj - b$ and $b > 0$.)

Then $j \neq 1$ and $j \neq b$. (Reason: j cannot be 1 because if $j = 1$ then it would follow from $b = jk$ that $k = b$, and we know that $k \neq b$. And j cannot be b because if $j = b$ then it would follow from $b = jk$ that $k = 1$, and we know that $k \neq 1$.)

Then $j < b$ and $k < b$. (Reason: $k \geq 1$, because $k \in \mathbb{N}$; so $k > 1$, because $k \neq 1$; so $k \geq 2$; and then if j was $\geq b$ it would follow that $jk \geq 2j > j > b$, but $jk = b$. The proof that $k < b$ is exactly the same.)

Hence $j \notin B$ (because b is the smallest member of B, and $j < b$). And $j \geq 2$ (because $j > 1$). This means that j is a product of primes (because if j wasn't a product of primes it would be in B).

Similarly, k is a product of primes. So we can write

$$j = \prod_{i=1}^{m} p_i \qquad \text{and} \qquad k = \prod_{\ell=1}^{\mu} q_\ell\,,$$

where $m \in \mathbb{N}$, $\mu \in \mathbb{N}$, and the p_i and the q_ℓ are primes. But then

$$b = \Big(\prod_{i=1}^{m} p_i\Big) \times \Big(\prod_{\ell=1}^{\mu} q_\ell\Big)\,,$$

so $\boxed{b \text{ is a product of primes}}$. (Precisely: define u_j, for $j \in \mathbb{N}$, $1 \leq j \leq m+\mu$, by the formula

$$u_j = \begin{cases} p_j & \text{if} \quad 1 \leq j \leq m \\ q_{j-m} & \text{if} \quad m+1 \leq j \leq m+\mu \end{cases}\,.$$

Then

$$b = \prod_{i=1}^{m+\mu} u_j\,.$$

And the u_j are prime, because each u_j is either one of the p_is or one of the q_ℓs.)

So $\boxed{b \text{ is a product of primes}}$.

But we know that $\boxed{b \text{ is not a product of primes}}$. So we got two contradictory statements.

This contradiction was derived by assuming that $B \neq \emptyset$. So $B = \emptyset$, and this proves that every natural number n such that $n \geq 2$ is a product of primes, which is our desired conclusion. **Q.E.D.**

16.6.3 The uniqueness question for the FTA

Remark 21. The ***fundamental theorem of arithmetic (FTA)*** says that every natural number greater than 2 can be written as a product of primes in a unique way. (That is, not only is the number equal to a product of primes, but there is only one way to write it as a product of primes.) Theorem 65 is a part of the FTA, namely, the assertion that if $n \in \mathbb{N}$ and $n \geq 2$ then n can be written as a product of primes.

What we have not proved is the uniqueness of the factorization. This is much more delicate, and we will prove it later.

At this point, just notice that even ***defining*** what "uniqueness" of the factorization of a natural number n into primes means is not a trivial question. For example, we can write the number 6 as a product of primes in this way:

$$6 = 2 \times 3\,,$$

but we can also write it as

$$6 = 3 \times 2\,.$$

Are these two expressions different ways of factoring 6 as a product of primes, or are they "the same"? Obviously, they must be "the same", because if they were different then the factorization of 6 as a product of primes would not be unique, and the FTA would not be true.

This means that we will have to be very precise, and define very carefully what "writing a number as a product of primes in a unique way" means. And this will be done later. □

17 The main theorems of elementary integer arithmetic II: the greatest common divisor of two integers and Bézout's lemma

Elementary integer arithmetic

Integer arithmetic is the study of the integers.
Elementary integer arithmetic is the study of the most basic facts about the integers. It is a body of theory that

- involves a number of important concepts, such as

  (**) divisibility,

  (**) prime numbers,

  (##) greatest common divisor,

- contains interesting and sometimes surprising results, such as

  (*#) the fundamental theorem of arithmetic,

  (##) Bézout's lemma,

  (##) Euclid's lemma,

  (**) Euclid's theorem on the existence of infinitely many prime numbers,

  and uses several powerful tools, such as

  (**) the principle of mathematical induction (PMI),

  (**) the well-ordering principle (WOP),

  (**) the division theorem.

(The items marked "(**)" have already been discussed in these notes. The items marked "(##)" will be discussed in this section. One item is marked "(#)", because we have already proved one half of it, whereas the other half has not yet been proved, but will be proved in this section.*

We now explain the concepts and results from the above list that have not been discussed yet, and prove the theorems.

17.1 The greatest common divisor of two integers

The first item in the list that is new to us is the concept of "greatest common divisor", so we begin by explainimg what this means.

In order to define "geatest common divisor",

1. We will first define "common divisor". This is going to be a *three-argument predicate* (because "c is a common divisor of a and b" is a statement about a, b and c that can be true or false depending on who a, b, c are).

2. Having defined "common divisor", the definition of "greatest common divisor" will just say the most obvious thing: a greatest common divisor of a and b is a common divisor that is the largest of all commmon divisors.

And here, finally, are the definitions:

The greatest common divisor of two integers

Definition 29. *Let a, b, g be integers. We say that c is a common divisor (or common factor) of a and b if c divides a and c divides b.* □

In other words,

$$c \text{ is a common divisor of } a \text{ and } b \iff (c|a \wedge c|b)\,. \qquad (17.379)$$

Definition 30. *Let a, b, g be integers. We say that g is a greatest common divisor of a and b if*

1. *g is a common divisor of a and b.*

2. *If c is any common divisor of a and b, then $c \leq g$.* □

In other words: ***a greatest common divisor of the integers a, b, is a common divisor that is greater that or equal to every common divisor of a and b.***

We are going to use "GCD" as an abbreviation for "greatest common divisor. Then

$$(\forall a \in \mathbb{Z})(\forall b \in \mathbb{Z})(\forall g \in \mathbb{Z})\Bigg(g \text{ is a GCD of } a \text{ and } b$$
$$\iff \Big(g|a \wedge g|b \wedge (\forall c \in \mathbb{Z})\big((c|a \wedge c|b) \implies c \leq g\big)\Big)\Bigg). \quad (17.380)$$

17.1.1 When do we use "a" and when do we use "the"?

*The question whether we can talk about **the** smallest member of a set, or we have to say **"a"** smallest member, was discussed in Section 16.1.2, on page 308. Here we deal with the same issue in the context of the greatest common divisor: can we talk about **the** greatest common divisor of a and b, or do we have to say **"a"** greatest common divisor of a and b?*

The general answer we gave in Section 16.1.2 was: *we talk about "**the** ZZZ" if there is only one ZZZ, and we talk about "**a** ZZZ" if there is more that one ZZZ.*

17.1.2 Uniqueness of the greatest common divisor

So which one is it? Shall we talk about "the" greatest common divisor of two integers, or about "a" greatest common divisor?

So far, in Definition 30, I talked about ***a*** greatest common divisor, because we didn't know yet if there is only one or more than one greatest common divisor of two given integers.

But now we are going to ***prove*** that the greatest common divisor, if it exists, is unique. And once we know that, we will be able to talk about ***the*** greatest common divisor of two integers.

Theorem 66. *Let a, b be integers. Then, if a greatest common divisor of a and b exists, it follows that a and b have only one greatest common divisor.*

Proof. To prove that there is only one GCD of a and b, we assume that g_1 and g_2 are GCDs of a and b, and prove that $g_1 = g_2$.

Since g_1 is a GCD of a and b, the definition of "GCD" tells us that $g_1|a$ and $g_1|b$.

Since g_2 is a GCD of a and b, the definition of "GCD" tells us that if c is any integer such that $c|a$ and $c|b$, then $c \leq g_2$. And we can apply this with g_1 in the role of c. Since $g_1|a$ and $g_1|b$, it follows that $g_1 \leq g_2$.

Exactly the same argument works to prove that $g_2 \leq g_1$.

Since $g_1 \leq g_2$ and $g_2 \leq g_1$, it follows that $g_1 = g_2$. **Q.E.D.**

So from now on we can talk about "***the*** GDC of a and b". And we can give it a name. So we shall call it "$GCD(a, b)$".

If a, b are integers, and the greatest common divisor of a and b exists, then "$GDC(a, b)$" is the name of the GCD of a and b.

Example 75.

1. $GCD(5, 7) = 1$. *Reason:* The only common divisors of 5 and 7 are 1 and -1. And 1 is the largest of the two, so $1 = GCD(5, 7)$.
2. $GCD(5, 15) = 5$. *Reason:* The common divisors of 5 and 15 are 1, -1, 5 and -5. And 5 is the largest of these four integers, so $5 = GCD(5, 15)$.
3. $GCD(18, 30) = 6$. *Reason:* The common divisors of 18 and 30 are 1, -1, 2, -2, 3, -3, 6, and -6. And 6 is the largest of these integers, so $6 = GCD(18, 30)$.
4. $GCD(28, 73) = 1$. *Reason:* 73 is prime. So the only factors of 73 are 1, -1, 73 and -73. But 73 and -73 are not factors of 28. So the only common divisors of 28 and 73 are 1 and -1. And 1 is the largest one. So $1 = GCD(28, 73)$.
5. $GCD(28, 0) = 28$. *Reason:* Every integer k is a factor of 0, because $0 = 0 \times k$, so $(\exists u \in \mathbb{Z})0 = uk$, so $k|0$. So the common factors of 28 and 0 are the factors of 28. And the largest of those factors is 28. So $28 = GCD(28, 0)$.
6. $GCD(-28, 0) = 28$. *Reason:* Every integer k is a factor of 0, as explained before. So the common factors of -28 and 0 are the factors of -28. And the largest of those factors is 28. So $28 = GCD(-28, 0)$.

In all the examples in the previous list, the GDC turned out to be positive. We can prove easily that this is a general fact:

Theorem 67. *Let a, b be integers such that the greatest common divisor $GCD(a, b)$ exists. Then*

$$GCD(a, b) \geq 1\,.$$

Proof. $GCD(a, b)$ is greater than or equal to every common factor of a and b. And 1 is a common factor of a and b. So $GCD(a, b) \geq 1$. **Q.E.D.**

17.2 Bézout's lemma

An extremely important, and rather surprining, fact about geratest common divisors is ***Bézout's lemma***.

Before I discuss the general statement of Bézout's lemma, let us look at an example.

17.3 Bézout's lemma: an example

Problem 89. Suppose you have two bottles. One of the bottles has a volume of exactly 500 milliliter and the other one has a volume of 700 milliliter. In addition, you have a large container and you can pour water from the bottles to the container or from the container to the bottles.

Show how, using these two bottles, you can end up with exactly 100 milliliter of water in the container.

Solution. The greatest common divisor of 500 and 700 is 100. By Bézout's Lemma, there exist integers u, v such that

$$100 = 500u + 700v\,. \tag{17.381}$$

Integers u, v for which (17.381) holds can be oomputed, for example, using the Euclidean algorithm. We find that $u = 3$, $v = -2$ are possible values[96] of u and v. So

$$100 = (-2) \times 700 + 3 \times 500\,.$$

So we can measure exactly 100 milliliters if water as follows:

- Fill the bottle whose volume is 500 milliliters with water, and then empty the bottle by pouring its contents into the large container. Do this three times.
- You will end up with 1500 milliliters in the container.
- Now pour water from the container into the bottle whose volume is 700 milliliters, until you fill it, and then empty the bottle. Do this twice. This will remove 1400 milliliters from the large container.
- So you will end up with 100 milliliters in the container, as desired

[96] But they are *not* the only possible values. Other values are, for example, $u = -4$, $v = 3$.

17.3.1 Bézout's lemma: the statement

And now, here is Bézout's Lemma.

Bézout's lemma

If a and b are two integers that are not both equal to zero, then

1. the GCD of a and b exists,
2. the GCD of a and b is equal to the sum of a multiple of a and a multiple of b. That is, there exist integers u, v such that

$$GCD(a, b) = ua + vb\,. \tag{17.382}$$

17.3.2 Integer linear combinations

Definition 31. *If a, b, c are integers, we say that c is an integer linear combination of a and b if there exist integers u, v such that*

$$c = ua + vb\,. \tag{17.383}$$

We are going to use "ILC" as an abbreviation for "integer linear combination".

Example 76.

1. 6 is an ILC of 9 and 33, because $6 = 8 \times 9 + (-2) \times 33$.
2. 1 is an ILC of 61 and 12, because $1 = 1 \times 61 + (-5) \times 12$.
3. 0 is an ILC of any two integers a, b, because $0 = b \times a + (-a) \times b$.
4. 3 is an ILC of -7 and 2, because $3 = 1 \times (-7) + 5]times2$.

Definition 32. *A positive integer linear combination of two integers a, b is an integer c such that $c > 0$ and c is an ILC of a and b.* □

17.3.3 A stronger version of Bézout's lemma

Bézout's lemma, stronger version

Theorem 68 *Let a, b be integers. Then:*

1. *If $a = 0$ and $b = 0$, then a greatest common divisor of a and b in the sense of Definition 30 does not exist.*
2. *If $a \neq 0$ or $b \neq 0$, then*
   (a) *The greatest common divisor $GCD(a, b)$ of a and b exists,*
   (b) *$GCD(a, b)$ is the smallest of all positive integers that are integer linear combinations of a and b.*

Example 77. Let $a = 12$, $b = 21$. Then the greatest common divisor of a and b is clearly 3.

Is 3 a positive ILC of a and b? Indeed,

$$3 = 2 \times 12 + (-1) \times 21\,,$$

so 3 is an ILC of a and b, and 3 is positive, so 3 is a positive ILC of a and b?

Are there and positive ILC's of a and b that are smaller than 3? Clearly, there are not, because any ILC of 12 and 21 must be divisible by 3, so a positive ILC of 12 and 21 must be at least equal to 3.

So 3 is the smallest positive ILC of a and b. □

17.3.4 Bézout's lemma: the proof

Proof of Theorem 68.

Let us start by giving a name to the set of all integers c such that c is an integer linear combination of a and b. Let us call this set "ILC(a, b)".

So the set ILC(a, b) is defined as follows:

$$\text{ILC}(a, b) = \{\, c \in \mathbb{Z} : (\exists u \in \mathbb{Z})(\exists v \in \mathbb{Z}) c = ua + bv \,\}\,. \tag{17.384}$$

And now that we have defined the set ILC(a, b), we can say "$c \in \text{ILC}(a, b)$" instead of "c is an integer linear combination of a and b".

First let us look at the case when $a = 0$ and $b = 0$. In this case, every integer is a common factor of a and b, because every integer divides 0. So there is no largest integer that is a common factor of a and b. That is, the GDC of a and b does not exist.

Now let us look at the case when $a \neq 0$ or $b \neq 0$. Let S be the set of all positive ILC's of a and b. That is, let

$$S = \{n \in \mathbb{N} : n \in \text{ILC}(a, b)\}\,.$$

It is clear that one of the four numbers $a, -a, b, -b$ must be positive. (If $a \neq 0$ then either $a > 0$ or $-a > 0$. If $b \neq 0$ then either $b > 0$ or $-b > 0$.) And all four numbers belong to ILC(a, b). So one of the four numbers belongs to ILC(a, b) and is positive. Hence S is a nonempty set of natural numbers. By the well-ordering principle, S has a smallest member. And, in addition, we know that the smallest member of a subset of $\mathbb{R}$, if it exists, is unique. So we can talk about ***the*** smallest member of S.

Let us give a name to this smallest member; let us call it g. So

$$g \in S \wedge (\forall n \in S)\, g \leq n\,. \tag{17.385}$$

We want to prove that

(*) g is the greatest common divisor of a and b.

In order to prove (*), the definition of "greatest common divisor" tells us that we have to prove the following two things:

(*1) g is a common divisor of a and b; that is,

$$g|a \wedge g|b\,. \tag{17.386}$$

(*2) g is the largest of all common divisors of a and b; that is,

$$(\forall c \in \mathbb{Z})\Big((c|a \wedge c|b) \Longrightarrow c \leq g\Big)\,. \tag{17.387}$$

Since $g \in \mathrm{ILC}(a,b)$, we can pick integers u, v such that

$$g = ua + vb\,. \tag{17.388}$$

*Proof of (*1).* Using the division theorem, we can divide a by g with a remainder r. That is, we can pick integers q, r such that

$$a = gq + r \text{ and } 0 \leq r < g\,. \tag{17.389}$$

(The division theorem says "$0 \leq r < |g|$". But in our case we know that $g \in \mathbb{N}$, so $|g| = g$.)

Then

$$\begin{aligned} r &= a - gq \\ &= a - (ua + vb)q \\ &= a - uqa - vqb \\ &= (1 - uq)a + (-vq)b\,. \end{aligned}$$

So

$$r \in \mathrm{ILC}(a,b)\,. \tag{17.390}$$

We know that $r \geq 0$. Let us prove that $r = 0$, by contradition.

> Assume that $r \neq 0$.
> Since $r \geq 0$, it follows that $r > 0$.
> So r is an integer and $r > 0$.
> Since $r \in \mathrm{ILC}(a,b)$, it follows that $r \in S$.
> In addition, (17.389) tells us that $r < g$.
> So g is not the smallest member of S, because r is a member of S and $r < g$.
> But g is the smallest member of S.
> Hence

g is the smallest member of S and g is not the smallest member of S,

> which is a contradiction.

So we have derived a contradiction from the assumption that $r \neq 0$. Hence $r = 0$. Since $r = 0$ and $a = gq + r$, we can conclude that $a = gq$. Therefore $g|a$.

The proof that $g|b$ is identical, and we omit it.

So $\boxed{g|a \wedge g|b}$, and this completes the proof of (*1).

*Proof of (*2).* We want to prove the universal sentence (17.387).

Let $c \in \mathbb{Z}$ be arbitrary.

Assume that $c|a \wedge c|b$.

Then we can pick integers j, k such that

$$a = cj \text{ and } b = ck\,.$$

Since $g = ua + vb$, we get

$$\begin{aligned} g &= ua + vb \\ &= ucj + vck \\ &= c(uj + vk)\,. \end{aligned}$$

Furthermore, $uj + vk$ is an integer, because u, v, j and k are integers.

Hence c divides g.

Our goal is to prove that $c \leq g$. And for that purpose we distinguish two cases: either $c \leq 0$ or $c > 0$.

Case 1: $c \leq 0$. In this case, the conclusion that $\boxed{c \leq g}$ is obvious, because $c \leq 0$ and $g > 0$, since $g \in \mathbb{N}$.

Case 2: $c > 0$. In this case, we have

$$g = \ell c\,,$$

where $\ell = uj + vk$. Then ℓ is an integer.

Then ℓ must be > 0. (Reason: if ℓ was ≤ 0 then ℓc would be ≤ 0, since $c > 0$. But $\ell c = g$, and $g > 0$. So ℓ cannot be ≤ 0. So $\ell > 0$.)

Since ℓ is an integer, and $\ell > 0$, it follows that ℓ is a natural number. Hence $\ell \geq 1$.

Since $\ell \geq 1$ and $\ell c = g$, it must be the case that $\boxed{c \leq g}$. (Reason: if $c > g$, then it would follow that $\ell c > g$, because $\ell c \geq c$—since $\ell \geq 1$—and $c > g$. But $\ell c = g$.)

So we have shown that $c \leq g$. And this completes our proof. **Q**.**E**.**D**.

17.4 The Euclidean Algorithm

Bézout's Lemma says that, if a, b are integers and are not both sero, then

(a) the gratest common divisor g of a and b can be written as an integer linear combination

$$g = ua + vb \tag{17.391}$$

of a and b,

(b) g is actually the smallest positive integer linear combination of a and b.

The ***Euclidean algorithm*** is a method for computing g and finding the coefficients u, v of the expression (17.391) of g as an integer linear combination of a and b.

17.4.1 Description of the algorithm for the computation of the greatest common divisor

We are given two integers a, b, and we want to find their greatest common divisor g. And, in addition, we may also want to find an expression of g as an integer linear combination of a and b.

We first observe that the greatest common divisor of a and b is the same as the greatest common divisor of $|a|$ and $|b|$. So we might as well asume that a and b aare nonnegative.

Second, if $a = b =$, the greatest common divisor does not exist. Sp we will assume that $\neq 0$ or $b \neq 0$.

Third, if $a > 0$ and $b = 0$, then $g = a$, and an expression of g as an integer linear combination of a and b is

$$g = a \times 1 + b \times 0\,.$$

So we have the results we want and there is no need to do any computations.

Similarly, if $a = 0$ and $b > 0$, then $g = b$, and an expression of g as an integer linear combination of a and b is

$$g = a \times 0 + b \times 1\,,$$

so again there is no need to do any computations.

Finally, if a and b are equal, then $g = a$ (or $g = b$), and an expression of g as an integer linear combination of a and b is

$$g = a \times 1 + b \times 0\,,$$

so again there is no need to do any computations.

So we are going to assume from now on that the integers a, b are positive and not equal. After relabeling them, if necessary, we assume that $a > b > 0$.

Here is how the alogorithm proceeds to find the greatest common divisor of a and b:

- We compute a sequence r_0, r_1, r_2, ..., r_k of positive integers as follows:
  - $r_0 = a$, $r_1 = b$, and then
  - if $r_1 \neq 0$, then we write[97]

    $$r_0 = r_1 q_2 + r_2\,, \text{ where } q_2 \in \mathbb{Z}\,,\ r_2 \in \mathbb{Z}\,,\ 0 \leq r_2 < r_1$$

    (that is, we divide r_0 by r_1, and let q_2 be the quotient and r_2 be the remainde of the division);
  - if $r_2 \neq 0$, then we write

    $$r_1 = r_2 q_3 + r_3\,, \text{ where } q_3 \in \mathbb{Z}\,,\ r_3 \in \mathbb{Z}\,,\ 0 \leq r_3 < r_2$$

    (that is, we divide r_1 by r_2, and let q_3 be the quotient and r_3 be the remainder of the division);
  - if $r_3 \neq 0$, then we write

    $$r_2 = r_3 q_4 + r_4\,, \text{ where } q_4 \in \mathbb{Z}\,,\ r_4 \in \mathbb{Z}\,,\ 0 \leq r_4 < r_3$$

    (that is, we divide r_2 by r_3, and let q_4 be the quotient and r_4 be the remained of the division);
  - as so on .....
  - once we have computed $r_0, r_1, \ldots, r_k$ and $q_2, \ldots, q_k$, if $r_k \neq 0$, then we write

    $$r_{k-1} = r_k q_{k+1} + r_{k+1}\,, \text{ where } q_{k+1} \in \mathbb{Z}\,,\ r_{k+1} \in \mathbb{Z}\,,\ 0 \leq r_{k+1} < r_3$$

    (that is, we divide r_{k-1} by r_k, and let q_{k+1} be the quotient and r_{k+1} be the remainder of the division);
  - as so on .....
- the first time we get to $r_{k+1} = 0$, the process stops.

[97] Naturally, this is possible because of the division theorem, which not only tells us that q_2 and r_2 exist, but also guarantess that they are unique.

- The reason that we necessarily have to get to $r_{k+1} = 0$ at some point is this: if the process went on for ever, we would be generating numbers r_0, r_1, r_2, r_3 that are always positive and in addition are decreasing (that is, $r_0 > r_1 > r_2 > r_3 > \cdots$, and $r_j > 0$ for every j). But this is not possible because of the well-ordering principle: let S be the set whose members are all the r_j that are > 0. Then S is a nonempty set of natural numbers. By the WOP, S has a smallest member s. But then $s = r_k$ for some k. And then r_{k+1} must be zero, because if r_{k+1} was $\neq 0$ then it would be > 0, so it would be a member of S smaller than r_k, contradicting the fact that r_k is the smallest member of S.

- Then r_k is the greatest common divisor of a and b.

17.4.2 Proof that the algorithm works to compute the greatest common divisor of a and b

Since $r_{k-1} = r_k q_{k+1} + r_{k+1}$, and $r_{k+1} = 0$, we have

$$r_{k-1} = r_k q_{k+1} \,,$$

so r_k divides r_{k-1}.

Since $r_{k-2} = r_{k-1} q_k + r_k$, and r_k divides r_{k-1}, it follows that r_k divides r_{k-2} as well.

Since $r_{k-3} = r_{k-2} q_{k-1} + r_{k-1}$, and r_k divides r_{k-1}, and r_{k-2}, it follows that r_k divides r_{k-3} as well.

Continuing in this way, we show that r_k divides r_{k-1}, r_{k-2}, ..., until eventually we find that r_k divides r_0 and r_1, that is, r_k divides a and b.

So r_k is a common divisor of a and b.

Now we need to prove that r_k is the greatest common divisor of a and b. For this purpose, we have to prove that if c is any common divisor of a and b then $c \leq r_k$.

So let $c \in \mathbb{Z}$ be a common divisor of a and b. Then c divides r_0 and c divides r_1.

Since $r_0 = r_1 q_2 + r_2$, we have $r_2 = r_0 - r_1 q_1$ and, since c divides r_0 and r_1, it follows that c divides r_2.

Since $r_1 = r_2 q_3 + r_3$, we have $r_3 = r_1 - r_2 q_3$ and, since c divides r_1 and r_2, it follows that c divides r_3.

Continuing in this way, we prove that c divides r_0, r_1, r_2, r_3, r_4, and so on, until we end up proving that c divides r_k.

Since c divides r_k, it follows that $c \leq r_k$. (Proof: if $c \leq 0$ then $c \leq r_k$, because $r_k > 0$. If $c > 0$, then c and r_k are both positive integers. Since

$c|r_k$, we may write $r_k = cm$, $m \in \mathbb{Z}$. But then $m > 0$, so $m \in \mathbb{N}$, and then $m \geq 1$. It follows that $r_k = mc \geq c$. So $c \leq r_k$.)

So we have proved that r_k satisfies the two conditions in the definition of "greatest common divisor of a and b": it divides both a and b, and it is $\geq c$ for every common divisor c of a and b.

Therefore r_k is the greatest common divisor of a and b. **Q.E.D.**

17.4.3 How the algorithm can be used to write the greatest common divisor as an integer linear combination of a and b

Having computed the greatest common divisor r_k of a and b, it turns out that, if we are interested, we can also use our computation to express r_k as an integer linear combination of a and b.

The key point is this: *whenever two integers u, v are integer linear combinations of a and b, it follows that every integer w which is an integer linear combination of u and v can be expressed as an integer linear combination of a and b.*

(This how this can be done: write

$$u = ma + nb\,,\ v = pa + qb\,,\ w = ru + sv\,,\ \ m, n, p, q, r, s \in \mathbb{Z}\,.$$

Then

$$\begin{aligned} w &= ru + sv \\ &= r(ma + nb) + s(pa + qb) \\ &= rma + rnb + spa + sqb \\ &= (rm + sp)a + (rn + sq)b\,, \end{aligned}$$

so $w = (rm + sp)a + (rn + sq)b$ is the desired expression of w as an integer linear combination of a and b.)

Using this, we can successively express r_0, r_1, r_2, r_3, ..., as integer linear combinations of a and b as follows:

- r_0 and r_1 are integer linear combinations of a and b, because $r_0 = a$ and $r_1 = b$;
- r_2 is an integer linear combination of r_0 and r_1, because $r_2 = r_0 - r_1 q_1$, so r_2 is an integer linear combination of a and b,
- r_3 is an integer linear combination of r_1 and r_2, because $r_3 = r_1 - r_2 q_2$; since r_1 and r_2 are integer linear combinations of a and b, it follows that r_3 is an integer linear combination of a and b,

- r_4 is an integer linear combination of r_2 and r_3, because $r_4 = r_2 - r_3 q_4$; since r_2 and r_3 are integer linear combinations of a and b, it follows that r_4 is an integer linear combination of a and b,
- continuing in this way, we end up finding an expression for r_k as an integer linear combination of a and b.

Example 78. Let us find the greatest common divisor of a and b, if $a = 700$, $b = 500$, using the Euclidean algorithm.

We let $r_0 = 700$, $r_1 = 500$. We then divide r_0 by r_1, and find q_2, r_2 such that $r_0 = r_1 q_2 + r_2$. We get

$$700 = 500 \times 1 + 200\,,$$

so $q_2 = 1$, $r_2 = 200$.

We then divide r_1 by r_2, and find q_3, r_3 such that $r_1 = r_2 q_3 + r_3$. We get

$$500 = 200 \times 2 + 100\,,$$

so $q_3 = 2$, $r_3 = 100$.

Next, we r_2 by r_3, and find q_4, r_4 such that $r_2 = r_3 q_4 + r_4$. We get

$$200 = 100 \times 2 + 0\,,$$

so $q_4 = 2$, $r_4 = -0$.

Since $r_4 = 0$, the process stops here, and the greatest common divisor is r_3, that is, 100.

To express the greatest common divisor as an integer linear combination of 700 and 500, we successively express r_0, r_1, r_2, r_3 as integer linear combinations of 700 and 500:

$$\begin{aligned} r_0 &= 700\,, \\ r_1 &= 500\,, \\ r_2 &= r_0 - r_1 q_2 \\ &= 700 - 500\,, \\ r_3 &= r_1 - r_2 q_3 \\ &= 500 - (700 - 500) \times 2 \\ &= 3 \times 500 + (-2) \times 700\,, \end{aligned}$$

so we end up with $\boxed{100 = 3 \times 500 + (-2) \times 700}$, which is the expression of the greatest common divisor 100 as an integer linear combination of 500 and 700 that we used in our solution of problem 89. □

Problem 90. ***Prove*** that if a, b are nonzero integers, g is the greatest common divisor of a and b, and $|a| > 1$ and $|b| > 1$, then g can be expressed as an integer linear combination

$$g = ua + vb\,,\ \ u \in \mathbb{Z}\,,\ v \in \mathbb{Z}\,,$$

in such a way that $|u| < |b|$ and $|v| < |a|$.

Here is an example: Take $a = 5$, $b = 3$. Then $g = 1$. We can write $g = 7 \times 3 + (-4) \times 5$, so we can take $u = 7$ and $v = -4$. But these numbers are too big. Since $7 = 5 + 2$, we have

$$\begin{aligned} g &= 7 \times 3 + (-4) \times 5 \\ &= (5+2) \times 3 + (-4) \times 5 \\ &= 2 \times 3 + (3 + (-4)) \times 5 \\ &= 2 \times 3 + (-1) \times 5\,, \end{aligned}$$

so we now hav $g = ua + vb$ with $|u| < |b|$ and $|v| < |a|$.

Your job is to prove that this method for making u and v smaller always works. □

18 The main theorems of elementary integer arithmetic III: Prime numbers, Euclid's lemma, coprime integers

18.1 The definition of "prime number"

We repeat the definition of "prime number", given in section 4, on page 20

Definition of "prime number". A prime number is a natural number p such that

I. $p > 1$,

II. p does not have any natural number factors other than 1 and p. □

And here is another way of saying the same thing, in case you do not want to talk about "factors".

Another version of the definition of "prime number". A prime number is a natural number p such that

I. $p > 1$,

II. There do not exist natural numbers j, k such that $j > 1$, $k > 1$, and $p = jk$. □

18.2 Euclid's lemma: an important application of Bézout's lemma

Euclid's lemma is one of the most important technical results in elementary integer arithmetic. For example, ***Euclid's lemma is the key fact needed to prove the missing half of the Fundamental Theorem of Arithmetic (FTA), that is, the uniqueness of the prime factorization.***

And, as you will see, the key fact that makes the proof of Euclid's lemma work is Bézout's lemma.

Euclid's lemma is about the following question:

Question 3. *Suppose an integer p divides the product ab of two integers a, b. Does it follow that p must divide a or p must divide b?* □

The answer is "no" if a, b and p are arbitrary integers.

Example 79. 6 divides 2×3 (because $6 = 2 \times 3$) but 6 doesn't divide 2 and 6 does not divide 3. □

But it turns out that the answer s "yes" if p is prime, and this is what Euclid's lemma says:

Theorem 69. ***(Euclid's lemma)*** *If a, b, p are integers, such that p is prime and p divides the product ab, then p divides a or p divides b.*

Proof. To prove that $p|a \vee p|b$, we prove[98] that $(\sim p|a) \Longrightarrow p|b$. i.e., that if p does not divide a then p divides b.

Assume that p does not divide a. Since p is prime, the only natural numbers that are factors of p are 1 and p. And p is not a factor of a, because we are assuming that p does not divide a.

Therefore the greatest common divisor of p and a is equal to 1.

It then follows from Bézout's lemma that 1 is equal to the sum of a multiple of p and a multiple of a. That is, we can pick integers u, v such that

$$1 = up + va\,.$$

On the other hand, since p divides ab, we may pick an integer k such that

$$ab = pk\,.$$

Then

$$\begin{aligned} b &= b \times 1 \\ &= b \times (up + va) \\ &= ubp + vab \\ &= ubp + vpk \\ &= (ub + vk)p\,, \end{aligned}$$

so p divides b. **Q.E.D.**

[98] Why do we do that? This is so because of Rule $\vee_{prove}$, the rule for proving "$\vee$" sentences: if, assuming $\sim A$, you prove B, then you can go to $A \vee B$. And the reason for Rule $\vee_{prove}$ is this: suppose we want to prove $A \vee B$. There are two possibilities: either A is true or A is not true. If A is true then $A \vee B$ is true, and we are done. If A is false then, since we know how to prove B assuming $\sim A$, B follows, so "$A \vee B$" is true in this case as well. Here is another way to see this: "$A \vee B$" is false if and only if both A and B are false. And the implication "$(\sim A) \Longrightarrow B$" is false only if and only if the premise is true and the conclusion is false, that is, if and only if A is false and B is false. So "$A \vee B$" is false if and only if "$(\sim A) \Longrightarrow B$" is false. So "$A \vee B$" is true if and only if "$(\sim A) \Longrightarrow B$" is true. So proving "$A \vee B$" amounts to the same thing as proving "$(\sim A) \Longrightarrow B$". And to prove "$(\sim A) \Longrightarrow B$" we assume $\sim A$ and prove B.

18.3 Coprime integers

Coprime integers were defined in section 4.5.2, on page 64.

Here we review the definition, and reformulate the concept of coprimeness using the gratest common divisor.

Suppose a, b are two integers. A common factor of a and b is an integer that divides both a and b.

Clearly, every integer is divisible by 1 and by -1. So 1 and -1 are common factors of a and b, no matter who the integers a and b are. Since 1 and -1 are always common factors, they are not very interesting common factors. We call them ***the trivial common factors*** of a and b, because it is a trivial fact that they are always common factors.

The truly interesting question is whether a and b have other, ***nontrivial*** common factors. Two integers that do not have nontrivial common factors are said to be ***coprime***.

Definition 33. *If a, b are integers, we say that a and b are coprime (or that "a is coprime with b", or that "b is coprime with a") if a and b have no nontrivial common factors (that is, if the only integers f such that $f|a$ and $f|b$ are 1 and -1).* □

It follows trivially from Definition 33 that

Corollary 1. *If a and b are integers, then a and b are coprime if and only if they are not both equal to zero and $GCD(a, b) = 1$.* □

We now introduce a symbol for coprimeness:

If a and b are integers, we write

$$a \perp b$$

for "a and b are coprime".
For example:

$$\begin{array}{rcrrcrrcr} 3 & \perp & 7 & -12 & \perp & 55 & 1 & \perp & 0 \\ \sim 22 & \perp & 14 & \sim 78 & \perp & -15 & \sim 49 & \perp & 77 \end{array}.$$

18.3.1 An extension of Euclid's lemma: if $p|ab$ and $p \perp a$ then $p|b$

In this section we look at the following question:

Question 4. *If*

1. *p, a, b are integers,*
2. *p divides ab,*
3. *p does not divide a,*

can we conclude that p must divide b?

Euclid's lemma tells us that the answer is "yes" if p is prime.

But if p is not prime the answer could be "no", as we showed in Example 79.

It turns out that, using exactly the same strategy—nased on Bézout's lemma—that we used to prove Euclid's lemma, we can extend Euclid's lemma by proving that the answer is "yes" not only when p is prime but also in some cases when p is not prime.

What is needed is that p ***and*** a ***should be coprime***. This will always be the case when p is prime, because when p is prime and p does not divide a it follows that p and a are coprime.

Theorem 70. *If*

- *a, b, p, are integers,*
- *p is coprime with a,*
- *p divides the product ab,*

then p divides b.

Proof. Since $p \perp a$, the greatest common divisor $GCD(p, a)$ is equal to 1.
Using Bézout's lemma, we can pick integers u, v such that

$$ua + vp = 1\,. \tag{18.392}$$

Then, if we multiply both sides of (18.392) by b, we get

$$uab + vpb = b\,.$$

Since p divides ab, we can pick an integer k such that

$$ab = kp\,.$$

Then

$$\begin{aligned} b &= uab + vpb \\ &= ukp + vpb \\ &= (uk + vb)p\,, \end{aligned}$$

so p divides b. **Q.E.D.**

18.3.2 An important application of the theorem of section

18.3.3 Why is Theorem 70 "an extension of Euclid's lemma"?

We said before that Theorem 70 is "an extension of Euclid's lemma". To see why this is so, let me show how, once you have Theorem 70, Euclid's lemma follows easily:

An easy derivation of Euclid's lemma from Theorem 70: Suppose p is prime and p divides the product ab of two integers a, b. We want to prove that $p|a$ or $p|b$. For this purpose, we assume that p does not divide a and try to prove that p divides b.

Since p is prime and p does not divide a, p is coprime with a. Then Theorem 70 tells us that p divides b, which is exactly what we want to prove in order to prove Euclid's Lemma. **Q.E.D.**

18.3.4 Another extension of Euclid's lemma: if an integer is coprime with two integers, then it is coprime with their product

In addition to providing an easy way to prove Euclid's lemma, Theorem 70 has another important consequence:

Theorem 71. *If a, b, p, are integers, and p is coprime with a and with b, then p is coprime with the product ab.*

Theorem 71 is easy to remember: it says that

$$\text{If } p \perp a \text{ and } p \perp b \text{ then } p \perp ab\,.$$

Proof of Theorem 71.

Assume that p is not coprime with ab. Then p and ab have a common factor m such that $m > 1$.

Since $m|p$, and $p \perp a$, m must be coprime with a as well. (Reason: any common factor of m and a would be a common factor of p and a, since $m|p$. Since p and a do not have nontrivial common factors, m and a cannot have nontrivial common factors either.)

On the other hand, m divides ab, because $m|p$ and $p|ab$.

So m divides ab and m is coprime with a. By Theorem 70, m divides b.

Hence $m|b$, $m|p$, and $m > 1$. Therefore p and b have a nontrivial common factor.

It follows that $\boxed{p \text{ and } b \text{ are not coprime}}$.

But $\boxed{p \text{ and } b \text{ are coprime}}$.

So we have reached a contradiction, and this was the result of assuming that p is not coprime with ab.

Hence p is coprime with ab. **Q.E.D.**

Why is Theorem 71 "an extension of Euclid's lemma"? The reason is, once again, that from Theorem 71 one can easily derive Euclid's lemma.

An easy derivation of Euclid's lemma from Theorem 71: Suppose p is prime and p divides the product ab of two integers a, b. We want to prove that $p|a \vee p|b$. For this purpose, we assume that it is not true that $p|a \vee p|b$. Then p does not divide a and p does not divide b. Since p is prime and p does not divide a, p is coprime with a. Since p is prime and p does not divide b, p is coprime with b. Then Theorem 71 tells us that $\boxed{p \text{ is coprime with } ab}$.

On the other hand, we are assuming that $p|ab$, so p and ab have a nontrivial common factor[99], namely, p. So $\boxed{p \text{ is not coprime with } ab}$.

So we have reached a contradiction, and this happened because we assumed that it is not true that $p|a \vee p|b$. Hence $\boxed{\boxed{p|a \vee p|b}}$. **Q.E.D.**

18.4 Uniqueness of the coprime representation of a rational number

In section 16.5 we defined "coprime representation" (cf. Definition 28 and proved that every rational number has a coprime representation (cf. Theorem 64). But we did not prove that the coprime representation of a rational number is unique. We now prove the uniqueness result.

Theorem 72. *If r is a rational number, then the coprime representation of r, whose existence was established in Theorem 64, is unique.*

Proof. Let r be an arbitrary rational number.

Assume that r has two coprime representations,

$$r = \frac{m}{n} \qquad \text{and} \qquad r = \frac{m'}{n'}\,, \tag{18.393}$$

where m, n, m', n' are integers, $n > 0$, $n' > 0$, $m \perp n$, and $m' \perp n'$.

We want to prove that $m = m'$ and $n = n'$.

It follows from 18.393 that

$$\frac{m}{n} = \frac{m'}{n'}\,. \tag{18.394}$$

[99] Why is p a ***nontrivial*** common factor? Because p is prime, so $p > 1$.

Then

$$mn' = m'n\,. \tag{18.395}$$

Then $n'|m'n$.

Since $n' \perp m'$, it follows from Theorem 70 that $n'|n$.

So we can pick an integer j such that $n = n'j$.

Similarly, we can pick an integer k such that $n' = nk$.

Then $n = n'j = (nk)j = n(kj)$, so $n = n(kj)$.

Since $n \neq 0$, it follows that $kj = 1$.

Then either $k = j = 1$ or $k = j = -1$.

But $n > 0$ and $n' > 0$, so $j = k = 1$.

Therefore $\boxed{n = n'}$.

And then (18.395) implies that $\boxed{m = m'}$. **Q.E.D.**

18.5 A general theorem on irrationality of square roots: an important application of Bézout's lemma and Euclid's lemma

After having proved that various numbers such as $\sqrt{2}$, $\sqrt{3}$, $\sqrt{5}$, $\sqrt{28}$, $\sqrt{\frac{2}{3}}$, $\sqrt{\frac{27}{31}}$ are irrational, can we prove once and for all a general theorem that will include all these cases? The answer is "yes", and here is the theorem. Notice that all the irrationality results about square roots that we proved before follow easily from this theorem. (For example: if $r = 2$, then $r = \frac{2}{1}$ and $2 \perp 1$, so Theorem 73 tells us that $\sqrt{r}$ is irrational, because 2 is not the square of an integer; similarly, if $r = \frac{2}{3}$, then Theorem 73 tells us that $\sqrt{r}$ is irrational, because $2 \perp 3$ and 2 and 3 are not squares of integers.)

Theorem 73. *Let r be a rational number written as a quotient*

$$r = \frac{m}{n}\,, \tag{18.396}$$

where m and n are coprime integers and $n > 0$. Then either

(1) there does not exist a rational number s such that $s^2 = r$.

or

(2) m and n are both squares of integers.

Proof of Theorem 73:

Let r be an arbitrary rational number. We want to prove that either (1) or (2) is true. Using Rule $\vee_{prove}$, we will assume that (1) is false, and prove that (2) is true.

Assume that (1) is false.

Then we can pick a rational number s such that $s^2 = r$.

Using Theorem 64 (on page 317), the number s has a coprime representation.

That is, we can pick integers p, q such that

i. $q > 0$,
ii. $p \perp q$,
iii. $s = \frac{p}{q}$.

Then $s^2 = \frac{p^2}{q^2}$, so

$$r = \frac{p^2}{q^2}\,. \tag{18.397}$$

Since $p \perp q$, Theorem 71 implies that $p \perp q^2$. And then, since $q^2 \perp p$, using Theorem 71 once more we find that

$$p^2 \perp q^2\,.$$

Furthermore, q^2 is clearly positive.

So (18.396) and (18.397) are coprime representations of r.

Using the uniqueness of the coprime representation (Theorem 72), we conclude that

$$m = p^2 \qquad \text{and} \qquad n = q^2\,. \tag{18.398}$$

So (2) holds.

So, assuming that (1) is false we proved (2). By Rule $\vee_{prove}$, it follows that (1)$\vee$(2) is true. **Q**.**E**.**D**.

Theorem 73 can be generalized to powers greater that 2. For example, we have:

Theorem 74. *Let r be a rational number written as a quotient*

$$r = \frac{m}{n}\,, \tag{18.399}$$

where m and n are coprime integers and $n > 0$. Then either

(1) there does not exist a rational number s such that $s^3 = r$.

or

(2) m and n are both cubes of integers (that is, $(\exists p \in \mathbb{Z})(\exists p \in \mathbb{Z})(m = p^3 \wedge n = q^3)$).

Proof of Theorem 73:

YOU DO IT.

Problem 91. ***Prove*** Theorem 74. □

And, even more generally, we have

Theorem 75. *Let r be a rational number written as a quotient*

$$r = \frac{m}{n}\,, \tag{18.400}$$

where m and n are coprime integers and $n > 0$, and let k be a natural number. Then either

(1) there does not exist a rational number s such that $s^k = r$.

or

(2) m and n are both k-th powers of integers (that is, $(\exists p \in \mathbb{Z})(\exists p \in \mathbb{Z})(m = p^k \wedge n = q^k)$).

Proof of Theorem 75:

YOU DO IT.

Problem 92. ***Prove*** Theorem 75. □

18.6 Divisibility properties of products of several integers

18.6.1 An important notational convention: the sets $\mathbb{N}_k$

In what follows we will be making lots of statements about "the natural numbers $1, 2, \ldots, k$", that is "all the natural numbers j such that $j \leq k$". So it will be convenient to give a name to the set of all such js.

THE SETS $\mathbb{N}_k$ (A.K.A. $\{1, 2, \ldots, k\}$)

The expression "$\mathbb{N}_k$" stands for the set of all natural numbers that are less than or equal to k. That is,

$$\mathbb{N}_k = \{n \in \mathbb{N} : n \leq k\}. \tag{18.401}$$

Another notation often used for this set is "$\{1, \ldots, k\}$", or "$\{1, 2, \ldots, k\}$".
We will use "$\mathbb{N}_k$" when k is a natural number, and also when $k = 0$. (So $\mathbb{N}_k$ makes sense when $k \in \mathbb{N} \cup \{0\}$.)
Naturally, for $n = 0$ the set defined by (18.401) has no members, because there are no natural numbers k such that $k \leq 0$. So

$$\mathbb{N}_0 = \emptyset. \tag{18.402}$$

For example:

$$\mathbb{N}_0 = \emptyset, \quad \mathbb{N}_1 = \{1\}, \quad \mathbb{N}_2 = \{1, 2\},$$
$$\mathbb{N}_3 = \{1, 2, 3\}, \quad \mathbb{N}_4 = \{1, 2, 3, 4\}, \quad \mathbb{N}_5 = \{1, 2, 3, 4, 5\}.$$

Then

$$j \in \mathbb{N}_k$$

is just another way of saying "$j \in \mathbb{N}$ and $j \leq k$".

18.6.2 The generalized Euclid lemma

Theorem 69 (that is, Euclid's lemma) tells us that if p is a prime and a, b are integers such that p is prime and p divides the product ab, then p divides a or p divides b.

The ***generalized Euclid lemma*** answers the following more general question:

Question 5. *What happens if instead of two integers a, b we have three integers a, b, c? Is it still true that if $p|abc$ then $p|a$ or $p|b$ or $p|c$?*

What if we have four integers a, b, c, d. Is it still true that if $p|abcd$ then $p|a$ or $p|b$ or $p|c$ or $p|d$? □

The answer is "yes", for three, four, or any number of integers, as shown by the following result.

Theorem 76. *Let k be a natural number, and let $p, a_1, a_2, \ldots, a_k$ be integers such that*

1. *p is a prime number,*
2. *p divides the product $\prod_{j=1}^{k} a_j$.*

Then p divides one of the factors. That is, $(\exists j \in \mathbb{N}_k) p|a_j$,

Proof. We will prove this by induction.

We want to prove

$$(\forall k \in \mathbb{N})(\forall p, a_1, a_2, \ldots, a_k \in \mathbb{Z}) \left(\left(p \text{ is a prime number} \wedge p \Big| \prod_{j=1}^{k} a_j \right) \Longrightarrow (\exists j \in \mathbb{N}_k) p|a_j \right). \quad (18.403)$$

Sentence (18.403) is a closed sentence. i.e., a sentence with no open variables, because the sentence contains the variables $k, p, a_1, a_2, \ldots, a_k$ and j, but they are all quantified, so no variables are open.

We can express sentence (18.403) as "$(\forall k \in \mathbb{N}) P(k)$", where $P(k)$ be the sentence

$$(\forall p, a_1, a_2, \ldots, a_k \in \mathbb{Z}) \left(\left(p \text{ is a prime number} \wedge p \Big| \prod_{j=1}^{k} a_j \right) \Longrightarrow (\exists j \in \mathbb{N}_k) p|a_j \right). \quad (18.404)$$

Then $P(k)$ is a sentence with one open variable, and the open variable is k. So $P(k)$ is exactly the kind of sentence for which we can expect to be able to prove "$(\forall k \in \mathbb{N})P(k)$" by induction.

Now let us prove "$(\forall k \in \mathbb{N})P(k)$" by induction.

Base step. We have to prove $P(1)$. But $P(1)$ says

$$(\forall p, a_1 \in \mathbb{Z})\left(\left(p \text{ is a prime number} \wedge p \Big| \prod_{j=1}^{1} a_j\right) \Longrightarrow (\exists j \in \mathbb{N}_1) p|a_j\right). \tag{18.405}$$

But $\mathbb{N}_1$ is just the set $\{1\}$, so "$(\exists j \in \mathbb{N}_1)p|a_j$" just amounts to saying "$p|a_1$".

Furthermore. $\prod_{j=1}^{1} a_j = a_1$. So $P(1)$ actually says

$$(\forall p, a_1 \in \mathbb{Z})\left(\left(p \text{ is a prime number} \wedge p|a_1\right) \Longrightarrow p|a_1\right). \tag{18.406}$$

And this is clearly true. So (18.406) is true.

Hence $P(1)$ is true.

Inductive step. We want to prove that

$$(\forall k \in \mathbb{N})(P(k) \Longrightarrow P(k+1)). \tag{18.407}$$

Let $k \in \mathbb{N}$ be arbitrary.

Assume that $P(k)$ is true.

Then

$$(\forall p, a_1, a_2, \ldots, a_k \in \mathbb{Z})$$
$$\left(\left(p \text{ is a prime number} \wedge p \Big| \prod_{j=1}^{k} a_j\right) \Longrightarrow (\exists j \in \mathbb{N}_k) p|a_j\right). \tag{18.408}$$

We want to prove $P(k+1)$, that is,

$$(\forall p, a_1, a_2, \ldots, a_k, a_{k+1} \in \mathbb{Z})$$
$$\left(\left(p \text{ is a prime number} \wedge p \Big| \prod_{j=1}^{k+1} a_j\right) \Longrightarrow (\exists j \in \mathbb{N}_{k+1}) p|a_j\right). \tag{18.409}$$

So let $p, a_1, a_2, \ldots, a_k, a_{k+1}$ be arbitrary integers such that

1. p is a prime number.

2. p divides $\prod_{j=1}^{k+1} a_j$.

We want to prove that $(\exists j \in \mathbb{N}_{k+1}) p|a_j$. i.e., that $p|a_j$ for some $j \in \mathbb{N}_{k+1}$.

The inductive definition of "$\prod$" tells us that

$$\prod_{j=1}^{k+1} a_j = \Big(\prod_{j=1}^{k} a_j\Big) a_{k+1}\,.$$

So

$$p\Big|\Big(\prod_{j=1}^{k} a_j\Big) a_{k+1}\,.$$

Euclid's lemma tells us, since p is prime, that if p divides a product uv of two integers then $p|u$ or $p|v$. In our case, if we take $u = \prod_{j=1}^{k} a_j$ and $v = a_{k+1}$, the lemma tells us that either

(i) p divides $\prod_{j=1}^{k} a_j$

or

(ii) p divides a_{k+1}..

We now see what happens in each of these two cases.

Case (i): Assume that p divides $\prod_{j=1}^{k} a_j$. Then we can use $P(k)$ and conclude that p divides one of the factors, that is, we can conlude that $(\exists j \in \mathbb{N}_k) p|a_j$. So we may pick j in $\mathbb{N}_k$ such that $p|a_j$. Then obviously $j \in \mathbb{N}_{k+1}$, so $\boxed{(\exists j \in \mathbb{N}_{k+1}) p|a_j}$.

Case (ii): Assume that p divides a_{k+1}. Then it is is also true that $\boxed{(\exists j \in \mathbb{N}_{k+1}) p|a_j}$.

So in both cases $(\exists j \in \mathbb{N}_{k+1}) p|a_j$, so we have established the conclusion that $\boxed{\boxed{(\exists j \in \mathbb{N}_{k+1}) p|a_j}}$.

We have proved this for arbitrary integers $p, a_1, a_2, \ldots, a_k, a_{k+1}$ such that p is a prime number and p divides $\prod_{j=1}^{k+1} a_j$.

Hence we have proved $P(k+1)$.

Since we have proved $P(k+1)$ assuming $P(k)$, we have proved the implication $P(k) \Longrightarrow P(k+1)$.

Since we have proved $P(k) \Longrightarrow P(k+1)$ for arbitray $k \in \mathbb{N}$, we have proved $(\forall k \in \mathbb{N})(P(k) \Longrightarrow P(k+1))$.

This completes the inductive step.

So we have proved $(\forall k \in \mathbb{N})P(k)$. **Q.E.D.**

18.6.3 A further extension of Euclid's lemma: if an integer p is coprime with several integers, then it is coprime with their product

Theorem 71 tells us that if an integer p is coprime with two integers a, b, then it is coprime with the product ab.

We now consider the following question:

Question 6. *What happens if instead of two integers a, b we have three integers a, b, c? Is it still true that if $p \perp a$, $p \perp b$, and $p \perp c$, then $p \perp abc$?*

What if we have four integers a, b, c, d. Is it still true that if $p \perp a$, $p \perp b$, $p \perp c$, and $p \perp d$, then $p \perp abcd$? □

The answer is "yes", for three, four, or any number of integers, as the following general theorem states.

Theorem 77. *Let n be a natural number, and let $p, a_1, a_2, \ldots, a_n$ be integers such that p is coprime with a_j for every $j \in \mathbb{N}_n$. Then p is coprime with the product $\prod_{j=1}^{n} a_j$.*

Proof. **YOU DO THIS.**

Problem 93. ***Prove*** Theorem 77 by induction. Use the inductive definition of $\prod_{j=1}^{n} a_j$, and use Theorem 71.

HINT: The proof is very similar to the proof of Theorem 76. I suggest that you read the proof of Theorem 76 carefully and use exactly the same pattern to prove Theorem 77. □

18.6.4 Another proof of the generalized Euclid lemma

Theorem 69 (that is, Euclid's lemma) tells us that If p is a prime and a, b, are integers such that p is prime and p divides the product ab, then p divides a or p divides b.

The ***generalized Euclid lemma*** answers the more general question "what happens if instead of two integers a, b we have three integers a, b, c? Or four integers a, b, c, d" Or, more generally, any number n of integers.

We answered this question by proving the generalized Euclid lemma (Theorem 76). Here I am giving you another proof of Theorem 76, based on Theorem 77.

Proof of Theorem 76 using Theorem 77.

Let $p, a_1, a_2, \ldots, a_k$ be integers such that p is prime and p divides $\prod_{j=1}^{k} a_j$. We want to prove that p divides one of the a_j.

Assume that $\boxed{p \text{ does not divide any of the } a_j}$.

Then, for each j, p is coprime with a_j. (Reason: since p is prime the only natural numbers that divide p are 1 and p. Since p does not divide a_j, the only natural number that divides both p and a_j is 1. So the greatest common divisor of p and a_j is 1. Then $p \perp a_j$.)

According to Theorem 77, it follows that $p \perp \prod_{j=1}^{k} a_j$.

But then $\boxed{p \text{ does not divide the product } \prod_{j=1}^{k} a_j}$.

But $\boxed{p \text{ divides the product } \prod_{j=1}^{k} a_j}$.

So we have reached a contradiction, by assuming that p does not divide any of the a_j. So $\boxed{p \text{ must divide one of the } a_j}$. **Q.E.D.**

18.7 Divisibility of an integer by the product of two or more integers

In this section we look at the following question:

Question 7. *If an integer q is divisible by several integers $a_1, a_2, \ldots, a_n$ when can we conclude that q is divisible by the product $a_1 \cdot a_2 \cdot \ldots \cdot a_n$?* □

It is clear that the answer is "not always".

Example 80. Let $a = 6$, $b = 4$, $q = 12$. Then 12 is divisible by a and by b, but it is clearly not divisible by ab, since $ab = 24$. □

18.7.1 Divisibility of an integer by the product of two integers

We now answer Question 7 for $n = 2$, i.e., for the case of two integers. is: ***if*** $a|q$ ***and*** $b|q$***, then we can conclude that*** q ***is divisible by the product*** ab ***if*** a ***and*** b ***are coprime***.

Indeed, we can prove:

Theorem 78. *If*

1. *a, b, q are integers,*
2. *a divides q and b divides q,*
3. *a and b are coprime,*

then ab divides q.

Proof. Since a and b are coprime, we may pick integers u, v such that $1 = ua + vb$.

Since q is divisible by a and by b, we can pick integers j, k such that $q = aj$ and $q = bk$. Then

$$\begin{aligned} q &= q \times 1 \\ &= q \times (ua + vb) \\ &= qua + qvb \\ &= (bk)ua + (aj)vb \\ &= ab(ku + jv)\,. \end{aligned}$$

Since $ku + jv$ is an integer, it follows that ab divides q. **Q.E.D.**

18.7.2 Divisibility of an integer by the product of several integers

Supose an integer n is divisible by three integers a, b, c. Can we conclude that n is divisible by the product abc?

What if n is divisible by four integers a, b, c, d? Can we conclude that n is divisible by the product $abcd$?

In general, let us look at the following question:

Question 8. *Suppose that*

1. *n is an integer,*
2. *k is a natural number,*
3. *$a_1, a_2, \ldots, a_k$ are integers,*
4. *n is divisible by all the a_j; that is, $(\forall j \in \mathbb{N}_k) a_j | n$.*

Can we conclude that the product $\prod_{j=1}^{k}$ divides n? □

For the case of two integers a_1, a_2, we know that the answer is "yes" if a_1 and a_2 are coprime. The answer for several integers $a_1, a_2, \ldots, a_k$ is similar: we have to require that $a_1, a_2, \ldots, a_k$ be ***pairwise coprime***. This means that $a_1 \perp a_2$, $a_1 \perp a_3$, $a_2 \perp a_3$, $a_1 \perp a_4$, $a_2 \perp a_4$, and so on. *Every pair a_i, a_j has to be coprime* (except of course when $i = j$; we do not want to demand, for example, that a_1 be coprime with a_1, because that would amount to requiring that a_1 be equal to 1). .

Definition 34. *Let $k \in \mathbb{N}$, and let $a_1, a_2, \ldots, a_k$ be integers. We say that $a_1, a_2, \ldots, a_k$ are pairwise coprime if for every $i \in \mathbb{N}_k$ and every $j \in \mathbb{N}_k$, if $i \neq j$ then a_i and a_j are coprime.* □

Theorem 79. *Assume that $n, a_1, a_2, \ldots, a_k$ are integers, k is a natural number, and*

1. *n is divisible by all the a_j; that is, $(\forall j \in \mathbb{N}_k) a_j | n$,*
2. *$a_1, a_2, \ldots, a_k$ are pairwise coprime, that is,*

$$a_i \perp a_j \qquad \text{whenever } i, j \in \mathbb{N}_k\,,\ i \neq j\,,$$

or, in more formal language, $(\forall i, j \in \mathbb{N}_k)(i \neq j \Longrightarrow a_i \perp a_j)$.

Then the product $\prod_{j=1}^{k} a_j$ divides n.

Proof. **YOU DO IT.**

Problem 94. ***Prove*** Theorem 79 by induction on k.

HINT: Let $P(k)$ be the statement

($\Diamond$) *If $n, a_1, a_2, \ldots, a_k$ are integers such that each a_j divides n, and the a_j are pairwise coprime, then the product $\prod_{j=1}^{k} a_j$ divides n,*

so, in formal language, $P(k)$ is

$$(\forall n, a_1, a_2, \ldots, a_k \in \mathbb{Z})$$
$$\left(\Big((\forall j \in \mathbb{N}_k) a_j | n \wedge (\forall i, j \in \mathbb{N}_k)(i \neq j \Longrightarrow a_i \perp a_j) \Big) \Longrightarrow \prod_{j=1}^{k} a_j | n \right) \tag{18.410}$$

Formula (18.410) contains the variables $n, i, j, k, a_1, a_2, \ldots, a_k$. But all these variables, except k, are quantified. So k is the only open variable. Hence *(18.410) is a one-variable predicate, and the open variable is k.* That's why we can call the predicate (18.410) $P(k)$, and you should prove by induction on k that $(\forall k \in \mathbb{N}) P(k)$.

In the inductive step of the proof, you should use Theorem 77 to conclude that a_{k+1} is coprime with $\prod_{j=1}^{k} a_j$, then use Theorem 78 to conclude, since $\prod_{j=1}^{k} a_j$ divides n, and a_{k+1} divides n, that $\prod_{j=1}^{k+1} a_j \Big| n$. $\square$

19 The main theorems of elementary integer arithmetic IV: The fundamental theorem of arithmetic

19.1 Introduction to the fundamental theorem of arithmetic

The *fundamental theorem of arithmetic* (FTA) says, roughly, that

(I) Every natural number n such that $n \geq 2$ is a product of prime numbers.

(II) The expression of n as a product of prime numbers is unique.

Statement (I) is an *existence* result: it says that

(E) For every $n \in \mathbb{N}$ such that $n \geq 2$ there exists a list

$$L = (p_1, p_2, \ldots, p_k)$$

such that $p_1, p_2, \ldots, p_k$ are prime numbers, and

$$n = \prod_{j=}^{k} p_j \,. \tag{19.411}$$

And we have already proved this, in Theorem 59.

The second half of the FTA is Statement (II), the *uniqueness* assertion: the list L such that (19.411) holds is unique.

We now have to prove (II). But before we do that, we have to make it precise. One possible meaning of (II) would be this:

(II_1) If $n \in \mathbb{N}$ and $n \geq 2$, then, if

$$L = (p_1, p_2, \ldots, p_k)$$

and

$$M = (q_1, q_2, \ldots, q_m)$$

are two lists of prime numbers such that

$$n = \prod_{j=1}^{k} p_j \quad \text{and} \quad n = \prod_{i=1}^{m} q_i \,, \tag{19.412}$$

then $L = M$. (That means "$m = k$, and $q_j = p_j$ for every $j \in \mathbb{N}_k$" , that is, $q_1 = p_1$, $q_2 = p_2$, ..., $q_k = p_k$.)

But it is easy to see that statement (II_1) cannot be true.

Example 81. Let $n = 6$, $p_1 = 2$, $p_2 = 3$, $q_1 = 3$, $q_2 = 2$. Then

$$6 = 2 \times 3 \text{ and } 6 = 3 \times 2\,,$$

so that

$$6 = p_1 p_2 \text{ and } 6 = q_1 q_2\,,$$

but it is not true that $p_1 = q_1$ and $p_2 = q_2$. □

In this example, it is clear what is really going on: ***it is not necessarily true that*** $p_1 = q_1$ ***and*** $p_2 = q_2$***. It could be the case that*** $p_1 = q_2$ ***and*** $p_2 = q_1$**.** In other words, "the p_js have to be the same as the q_js, but not necessarily in the same order".

How can we say this precisely? Let us try a second option:

(II_2) If $n \in \mathbb{N}$ and $n \geq 2$, then, if

$$L = (p_1, p_2, \ldots, p_k)$$

and

$$M = (q_1, q_2, \ldots, q_m)$$

are two lists of prime numbers such that

$$n = \prod_{j=1}^{k} p_j \quad \text{and} \quad n = \prod_{j=1}^{m} q_j\,, \tag{19.413}$$

then $m = k$ and the set P whose members are the p_j; that is, the set

$$P = \{p \in \mathbb{N} : (\exists j \in \mathbb{N}_k) p = p_j\}\,, \tag{19.414}$$

is the same as the set Q whose members are the q_j, that is, the set

$$Q = \{q \in \mathbb{N} : (\exists j \in \mathbb{N}_m) q = q_j\}\,. \tag{19.415}$$

But it is easy to see that this cannot be the right formulation either.

Example 82. Let

$$n = 72\,, \text{ that is } n = 2 \times 2 \times 2 \times 3 \times 3\,. \tag{19.416}$$

Then Formula (19.416) gives us a factorization of n as product of primes, namely,

$$n = p_1 p_2 p_3 p_4 p_5\,, \quad \text{where } p_1 = 2\,,\ p_2 = 2\,,\ p_3 = 2\,,\ p_4 = 3\,,\ p_5 = 3\,.$$

We would like to say that, if we have any other factorization

$$n = q_1 q_2 \cdots q_m\,,$$

then the q_js must be "the same" as the p_js, meaning first of all, that $m = 5$, and second, that three of the q_js must be equal to 2, and two of the q_js must be equal to 3.

And just saying that the set of the p_j is the same as the set of the q_j is not enough. The set P defined by Equation (19.414) is just the set $\{2, 3\}$, i.e., the set whose members are 2 and 3. (Remember that, for a set P, an object p is a member of P or is not a member of P; there is no such thing as "being a member of P twice", or "being a member of P three times".)

We want the q_js to be "the same" as the p_js not just in the set sense (that is, the set Q is also the set $\{2, 3\}$), but in the much stronger sense that "there are five q_js; three of them are 2s and two of them are 3s". And Formulation (II_2) does not capture that. □

So, how shall we say what we want to say? Let us go back to our examples.

Example 83. For the factorization

$$6 = p_1 p_2 \text{ where } p_1 = 2 \text{ and } p_2 = 3\,,$$

we want to say that if $q_1, q_2, \cdots, q_m$ are primes and $6 = q_1 q_2 \cdots q_m$, then

- m must be 2, so the equation "$6 = q_1 q_2 \cdots q_m$"becomes "$6 = q_1 q_2$".
- q_1 must be 2 and q_2 must be 3.

We can achieve this if we limit ourselves to ***ordered factorizations*** of 6, i.e., factorizations of 6 in which 6 is expressed as a product $q_1 q_2 \cdots q_m$ of primes, but the q_j are required to be in ***increasing order***, that is, to be such that $q_1 \le q_2 \le q_3 \le \cdots \le q_m$. This excludes the factorization $6 = 3 \times 2$, and leaves $6 = 2 \times 3$ as the only possible prime factorization of 6. □

Example 84. For the factorization

$$72 = p_1 p_2 p_3 p_4 p_5 \text{ where } p_1 = 2\,,\ p_2 = 2\,,\ p_3 = 2\,,\ p_4 = 3\,,\ p_5 = 3\,,$$

we want to say that if $q_1, q_2, \cdots, q_m$ are primes and $72 = q_1 q_2 \cdots q_m$, then m must be 5, three of the q_j must be 2, and two of the q_j must be 3. Again, we can achieve that if we limit ourselves to ***ordered factorizations*** of 72, i.e., factorizations of 72 in which 72 is expressed as a product $q_1 q_2 \cdots q_m$ of primes, but the q_j are required to be in increasing order, that is, to be such that $q_1 \leq q_2 \leq q_3 \leq \cdots \leq q_m$. This excludes other factorizations such as $72 = 3 \times 3 \times 2 \times 2 \times 2$, or $72 = 3 \times 2 \times 2 \times 3 \times 2$, and leaves $72 = 2 \times 2 \times 2 \times 3 \times 3$ as the only possible prime factorization of 72. □

Examples 83 and 84 show us the path: we have to define "ordered factorization" precisely, and then the statement of the FTA will be: *every natural number n such that $n \geq 2$ has a unique ordered factorization as a product of prime numbers.*

19.1.1 Precise statetement of the fundamental theorem of arithmetic

19.1.2 Is a prime factorization a set of primes?

If we are going to say that "every natural number n such that $n \geq 2$ has a unique prime factorization", then, to begin with, we have to answer the following question:

Question 9. *What do we mean, exactly, by a **prime factorization** of an integer n?* □

A prime factorization is, of course, something like "several primes that multiplied together result in n".

But such vague language will not do. We have to give a precise definition.

1. First of all, "prime factorization" is not an entity[100], like water, or politics. We can say things like

   > Water is a transparent and nearly colorless chemical substance

   or

   > Politics is the process of achieving and exercising positions of governance or organized control over a human community, particularly a state.

[100] According to the Merriam-Webster dictionary, an entity is "something that has separate and distinct existence and objective or conceptual reality".

But we cannot say "prime factorization is ...".

2. "Prime factorization" is like "subset", or "factor", or "divisible", or "absolute value": it is a ***relational concept***, it has arguments:
   (a) You cannot say "factor is ...", because "factor", by itself, is not something that can be or not be anything.
   (b) But you can say things like "a is a factor of b".
   (c) You cannot say "divisible is ..." (or, even worse, "divisible is when ..."), because "divisible", by itself, is not something that can be or not be anything.
   (d) But you can say things like "a is divisible by b".
   (e) You cannot say "absolute value is ...", because "absolute value", by itself, is not something that can be or not be anything.
   (f) But you can talk about "the absolute value of x".
3. More precisely, "prime factorization" is a ***two-argument predicate***: we say things like "$\mathbf{P}$ is a prime factorization of n". The arguments are n and $\mathbf{P}$. And, clearly, n must be a number.
4. And we haven't yet answered the question ***what kind of a thing shall $\mathbf{P}$ be?***
5. A prime factorization $\mathbf{P}$ should be a single object, not "several things".
6. And we have seen that it is not a good idea to think of a prime factorization as a ***set*** of primes, because, for example, the factorization of 72 given by $72 = 2 \times 2 \times 2 \times 3 \times 3$ contains more information than the set $\{2, 3\}$. It contains the fact that 2 "occurs three times", and 3 "occurs twice".

The conclusion of all this is that *a "prime factorization" should not be a **set**: it should be a **finite list**.*

And, to make this precise, we need to say a few words about finite lists.

19.2 Finite lists

In this section we will use the sets $\mathbb{N}_k$. The meaning of "$\mathbb{N}_k$" is explained in section 18.6.1, on page 350.

Definition 35. *Let n be a natural number.*

1. *A finite list of length n consists of the specification, for each natural number j in the set $\mathbb{N}_n$, of an object a_j.*

2. *The a_j are called the entries of the list:*

   (a) *a_1 is the first entry,*

   (b) *a_2 is the second entry,*

   (c) *a_3 is the third entry,*

   and so on, so that, for example, a_{283} is the 283rd entry.

3. *The entries a_j of a finite list $\mathbf{a}$ could be numbers of any kind (integers, real numbers, complex numbers, integers modulo 37), or matrices, or aets, orpoints, or lines, or planes, or functions, or lists, or planets, or animals, or people, or books, or viruses, or mice, or atoms, or ghosts, or unicorns, or angels, objects of any kind whatsoever, concrete or abstract, real or imaginary.*

4. *Actually, the entries of a list do not all have to be objects of the same kind (whatever "pf the same kind" means). So for or example, you can perfectly well have a finite list $\mathbf{a} = (a_1, a_2, a_3, a_4, a_5)$ in which a_1 is the number 5, a_2 is Mickey Mouse, a_3 is Abraham Lincoln, a_4 is the word "cow", and a_5 is the Pacific Ocean.*

Remark 22. *There are finite lists and infinite lists. In this section, we will only be talking about finite lists. But infinite lists are very important, and we will come back to them later.* □

19.2.1 How to introduce, specify, and name lists

- In principle, any symbol or string of symbols can be used as the name of a list, so we could name a list "a", or "q", or "Alice", or "list-of-primes".

- But in these notes we will use **boldface lower-case letters** for lists.

- And often, when we use a boldface letter such as $\mathbf{a}$ or $\mathbf{b}$ or $\mathbf{p}$ or $\mathbf{x}$ for a list, we will use the same letter in *italic*, with a subscript, as the name of an entry of a list.

- So, for example, if $\mathbf{p}$ is a list, then we may write "p_1" for the first entry of $\mathbf{p}$, "p_2" for the second entry, and, in general, "p_j" for the j-th entry.

- So, if $\mathbf{p}$ is a list of length n, then p_j will make sense for every $j \in \mathbb{N}_n$.
- We will write
$$\mathbf{a} = (a_j)_{j=1}^n \text{ or } \mathbf{a} = (a_j)_{k \in \mathbb{N}_n} \tag{19.417}$$
to indicate that $\mathbf{a}$ *is a finite list or length* n *and, for each* $j \in \mathbb{N}_n$, *the* j*-th entry of* $\mathbf{a}$ *is called* a_j.
- For short lists weill write (a_1), or (a_1, a_2), or (a_1, a_2, a_3), or (a_1, a_2, a_3, a_4), rather than $(a_j)_{j=1}^1$, or $(a_j)_{j=1}^2$, or $(a_j)_{j=1}^3$, or $(a_j)_{j=1}^4$.

And here are some examples of list specification:

Example 85. Suppose, for example, that we want to create a list of length 3, whose entries are the first three prime numbers, and we want to call it $\mathbf{a}$. We could write any of the following things to specify such a list:

$$\text{Let } \mathbf{a} = (2, 3, 5)\,, \tag{19.418}$$
$$\text{Let } \mathbf{a} = (a_1, a_2, a_3)\,, \text{ where } a_1 = 2,\, a_2 = 3,\, a_3 = 5\,, \tag{19.419}$$
$$\text{Let } \mathbf{a} = (a_j)_{j=1}^3\,, \text{ where } a_1 = 2,\, a_2 = 3,\, a_3 = 5\,, \tag{19.420}$$
$$\text{Let } \mathbf{a} = (a_j)_{j=1}^3\,, \text{ where } a_j \text{ is the } j\text{–th prime for } j = 1, 2, 3\,, \tag{19.421}$$
$$\text{Let } \mathbf{a} = (a_j)_{j=1}^3\,, \text{ where } a_j \text{ is the } j\text{–th prime for } j \in \mathbb{N}_3\,, \tag{19.422}$$
$$\text{Let } \mathbf{a} = (a_j)_{j=1}^3\,, \text{ where } (\forall j \in \mathbb{N}_3) a_j \text{ is the } j\text{–th prime}\,. \tag{19.423}$$

Example 86. Suppose we want to introduce the list of the first 500 prime numbers and give it a name. In this case, if we try to write something like (19.418) or (19.419) or (19.420) or (19.421) the formulas would get too long. But we can write

$$\text{Let } \mathbf{a} = (a_j)_{j=1}^{500}\,, \text{ where } a_j \text{ is the } j\text{–th prime for } j \in \mathbb{N}_{500}\,, \tag{19.424}$$
$$\text{Let } \mathbf{a} = (a_j)_{j=1}^{500}\,, \text{ where } (\forall j \in \mathbb{N}_{500}) a_j \text{ is the } j\text{–th prime}\,. \tag{19.425}$$

Example 87. Suppose we want to introduce the list of the first 500 squares of natural numbers and give it a name. In this case we can write one of the following:

$$\text{Let } \mathbf{a} = (a_j)_{j=1}^{500}\,, \text{ where } a_j \text{ is the } j\text{–th square for } j \in \mathbb{N}_{500}\,, \tag{19.426}$$
$$\text{Let } \mathbf{a} = (a_j)_{j=1}^{500}\,, \text{ where } (\forall j \in \mathbb{N}_{500}) a_j \text{ is the } j\text{–th square}\,, \tag{19.427}$$

but, since we have the *formula* $a_j = j^2$ for a_j, we have the additional options of writing one of the following:

$$\text{Let } \mathbf{a} = (a_j)_{j=1}^{500}\,, \text{ where } a_j = j^2 \text{ for } j \in \mathbb{N}_{500}\,, \tag{19.428}$$
$$\text{Let } \mathbf{a} = (j^2)_{j=1}^{500}\,. \tag{19.429}$$

Example 88. Suppose we want to introduce the list of all the U.S. presidents from George Washington to Donald Trump, in *chronological order*, that is, starting with George Washington and ending with Donald J. Trump.

We could do this by writing

$$\text{Let } \mathbf{a} = (a_{46-j})_{j=1}^{45}\,, \text{ where, for } j \in \mathbb{N}_{45}\,,\ a_j \text{ is the } j\text{−th president}\,. \quad \square$$

Now suppose we don't know how many presidents there have been from Washington to Trump, and we don't know that Trump is the 45-th president. We could write:

Let $\mathbf{a} = (a_j)_{j=1}^{N}$, where :

(a) N is the number of U.S. presidents from G. Washington to D. Trump and

(b) for $j \in \mathbb{N}_N$, a_j is the j−th U.S. president.

Example 89. Suppose we want to introduce the list of all the U.S. presidents from George Washington to Donald Trump, in *reverse chronological order*, that is, starting with George Washington and ending with Donald J. Trump.

We could do this by writing

$$\text{Let } \mathbf{a} = (a_j)_{j=1}^{45}\,, \text{ where, for } j \in \mathbb{N}_{45}\,,\ a_j \text{ is the } N{+}1{-}j\text{−th U.S. president}\,.$$

Remark 23. Often, one writes

$$\mathbf{a} = (a_1, \ldots, a_n)\,,$$

or

$$\mathbf{a} = (a_1, a_2, \ldots, a_n)\,,$$

instead of $\mathbf{a} = (a_j)_{j=1}^{n}$. I strongly prefer the $(a_j)_{j=1}^{n}$ notation, but I will accept the other one. $\square$

Remark 24. Pay attention to the following:

SETS VS. LISTS

1. Sets have ***members***, not entries.
2. Finite lists have ***entries***, not members.
3. In the set notation, we use ***braces***, as in "the set $\{x \in \mathbb{R} : x > 0\}$", or "the set $\{1, 2, 3, 4\}''$.
4. In the finite list notation, we use ***parentheses***, as in "the list $(p_j)_{j=1}^n$", or "the list $(2, 3, 5)$".
5. In a set S, an object a either is a member or is not a member. There is no such thing as "being a member of the set S twice".
6. In a finite list $\mathbf{a} = (a_j)_{j=1}^n$ it is possible for an object a to be the first entry of $\mathbf{a}$ (that is $a = a_1$) and also the second entry (that is, $a = a_2$) and the 25th entry (that is, $a = a_{25}$).
7. So ***a finite list can have repeated entries***, but ***a set cannot have repeated members***.

and to the following:

8. If $\mathbf{a}$ is a finite list, then we can associate to $\mathbf{a}$ a set $\mathrm{Set}(\mathbf{a})$, called the ***set of entries*** of the list $\mathbf{a}$

9. The set of entries of the list $\mathbf{a} = (a_j)_{j=1}^n$ is the set $\mathrm{Set}(\mathbf{a}$ given by

$$\boxed{\mathrm{Set}(\mathbf{a}) = \{x : (\exists j \in \mathbb{N}_n) x = a_j\}\,.}$$

This set is a totally different object from the list $\mathbf{a}$.

Remark 25. Not all books and journals use the same notation. So if you are reading a mathematics book or article you have to make sure to check which notations are being used. For example, some books use braces for lists, so they would write "the list $\{p_j\}_{j=1}^n$". I strongly prefer the parenthesis notation, and in this course this is the official notation, so we write "the list $(2, 2, 3, 4)$", or "the list $\mathbf{p} = (p_j)_{j=1}^n$", which are very different from "the set $\{2, 2, 3, 4\}$", or "the set $\{p : (\exists j \in \mathbb{N}_n) p = p_j\}$". (For example: the list $(2, 2, 3, 4)$ has four entries, but the set $\{2, 2, 3, 4\}$ has three members.) □

19.2.2 Equality of lists

We know that two sets A, B are equal if they have the same members. That is

$$A = B \iff (\forall x)(x \in A \iff x \in B)\,.$$

When are two finite lists equal?

Here is the asnwer:

Two lists

$$\mathbf{p} = (p_j)_{j=1}^n\,, \qquad \mathbf{q} = (q_j)_{j=1}^m\,,$$

are ***equal*** if

1. $n = m$,

and

2. $p_j = q_j$ for every $j \in \mathbb{N}_n$. (That is, $(\forall j \in \mathbb{N}_n) p_j = q_j$.)

Example 90. The lists $\mathbf{p} = (2, 2, 3)$ and $\mathbf{q} = (3, 2, 2)$ are *not* equal because, for example, the first entry of the first list is not equal to the first entry of the second list.

But, of course, the sets $\{2, 2, 3\}$ and $\{3, 2, 2\}$ are equal, because they are both equal to the set $\{2, 3\}$. □

Example 91. Let $\mathbf{P} = (p_j)_{j=1}^{45}$ be the list of all U.S. presidents from George Washington to Donald Trump. Then, for each $j \in \mathbb{N}_{45}$, p_j stands for "the j-th president of the United States".

Then $\mathbf{P}$ has 45 entries. Let S be the associated set $\mathrm{Set}(\mathbf{P})$. Then S is the set of all U.S. presidents from George Washington to Donald Trump. That is,

$$S = \{x : (\exists j \in \mathbb{N}_{45}) x = p_j\}\,.$$

How many members does S have?

If you guessed "45", you are wrong!

The correct answer is 44.

The reason for this is that Grover Cleveland was U.S. president from 1885 to 1889, and then again from 1893 to 1897. During his first presidency, he was the 22nd president. Then Benjamin Harrison served as the 23rd president, from 1889 to 1893, and after that Grover Cleveland was elected president again, and Congress decided that he would be counted at the 24th president, in addition to being counted as the 22nd president.

So the list $\mathbf{P}$ has a repeated entry: p_{22} is the same as p_{24}. The set $\mathrm{Set}(\mathbf{P})$ does not know this, because all a set knows is whether something (or somebody) is a member or not. So the set $\mathrm{Set}(\mathbf{P})$ has only 44 members. □

19.2.3 The sum, the product and the maximum and minimum of a finite list of real numbers

If $\mathbf{a}$ is a finite list of real numbers, then we can define several numbers associated to $\mathbf{a}$, using inductive definitions:.

Specifically, we will define

1. the ***sum*** $\sum \mathbf{a}$ of the entries of $\mathbf{a}$,
2. the ***product*** $\prod \mathbf{a}$ of the entries of $\mathbf{a}$,
3. the ***maximum*** $\operatorname{Max} \mathbf{a}$ of the entries of $\mathbf{a}$.
4. the ***minimum*** $\operatorname{Min} \mathbf{a}$ of the entries of $\mathbf{a}$.

In each of the cases, we start from a ***binary operation*** on $\mathbb{R}$, that is, an operation that can be performed on ***two*** real numbers, and extend it to finite lists.

The sum $\sum \mathbf{a}$ will be defined starting with the ***addition*** operation, i.e., the operation that for two real numbers x, y produces the number $x + y$.

The product $\prod \mathbf{a}$ will be defined starting with the ***multiplication*** operation, i.e., the operation that for two real numbers x, y produces the number $x \cdot y$.

The maximum $\operatorname{Max} \mathbf{a}$ will be defined starting with the ***maximum*** operation, i.e., the operation that for two real numbers x, y produces the number $\max(x, y)$ (the "maximum of a and b") defined as follows:

$$\max(x, y) = \begin{cases} x & \text{if} \quad x \geq y \\ y & \text{if} \quad y \geq x \end{cases} . \tag{19.430}$$

The minimum $\operatorname{Min} \mathbf{a}$ will be defined starting with the ***minimum*** operation, i.e., the operation that for two real numbers x, y produces the number $\min(x, y)$ (the "minimum of a and b") defined as follows:

$$\min(x, y) = \begin{cases} y & \text{if} \quad x \geq y \\ x & \text{if} \quad y \geq x \end{cases} . \tag{19.431}$$

Problem 95. The absolute value of a real number is defined as follows: if $x \in \mathbb{R}$, then the absolute value of x is the number $|x|$ given by

$$|x| = \begin{cases} x & \text{if} \quad x \geq 0 \\ -x & \text{if} \quad x \leq 0 \end{cases} . \tag{19.432}$$

Prove that

$$(\forall x \in \mathbb{R})(\forall y \in \mathbb{R})\, \max(x,y) = \frac{x+y+|x-y|}{2}$$

and

$$(\forall x \in \mathbb{R})(\forall y \in \mathbb{R})\, \min(x,y) = \frac{x+y-|x-y|}{2}\,.$$

The four operations $\sum$, $\prod$, Max, Min are defined as follows:

Definition 36. *Let* $\mathbf{a} = (a_j)_{j=1}^n$ *be a finite list of real numbers.*

1. *The* <u>sum</u> $\sum \mathbf{a}$, *or* $\sum_{j=1}^n a_j$, *is defined inductively as follows:*

$$\sum_{j=1}^{0} a_j = 0\,, \tag{19.433}$$

$$\sum_{j=1}^{1} a_j = a_1\,, \tag{19.434}$$

$$\sum_{j=1}^{n+1} a_j = \Big(\sum_{j=1}^{n} a_j\Big) + a_{n+1} \qquad \text{if } n \in \mathbb{N}\,. \tag{19.435}$$

2. *The* <u>product</u> $\prod \mathbf{a}$, *or* $\prod_{j=1}^n a_j$, *is defined inductively as follows:*

$$\prod_{j=1}^{0} a_j = 1\,, \tag{19.436}$$

$$\prod_{j=1}^{1} a_j = a_1\,, \tag{19.437}$$

$$\prod_{j=1}^{n+1} a_j = \Big(\prod_{j=1}^{n} a_j\Big) \times a_{n+1} \qquad \text{if } n \in \mathbb{N}\,, \tag{19.438}$$

$$\tag{19.439}$$

3. *The* <u>maximum</u> $\operatorname{Max} \mathbf{a}$, *or* $\operatorname{Max}_{j=1}^n a_j$, *is defined inductively as follows:*

$$\operatorname{Max}_{j=1}^{1} a_j = a_1\,, \tag{19.440}$$

$$\operatorname{Max}_{j=1}^{n+1} a_j = \max\Big(\operatorname{Max}_{j=1}^{n} a_j, a_{n+1}\Big) \qquad \text{if } n \in \mathbb{N}\,. \tag{19.441}$$

4. The <u>minimum</u> $\operatorname{Min}\mathbf{a}$*, or* $\operatorname{Min}_{j=1}^{n} a_j$*, is defined inductively as follows:*

$$\operatorname{Min}_{j=1}^{1} a_j = a_1\,, \tag{19.442}$$

$$\operatorname{Min}_{j=1}^{n+1} a_j = \max\left(\operatorname{Min}_{j=1}^{n} a_j, a_{n+1}\right) \qquad \text{if } n \in \mathbb{N}\,. \tag{19.443}$$

There are several facts about these operations that are fairly obvious, and whose proofs are very easy but very boring. I would urge you to practice by doing a few of these proofs, just to make sure that you can do them if you are asked to. Naturally, since the operations are defined inductively, the proofs will have to be by induction.

Before I tell you what these obvious facts are, let me define the ***concatenation*** of two lists: Roughly, the concatenation $\mathbf{a}\#\mathbf{b}$ is the list obtained by listing the entries of $\mathbf{a}$ first, and then the entries of $\mathbf{b}$.

Example 92.

1. Let

$$\mathbf{a} = (3,6,1,3,5)\,,$$
$$\mathbf{b} = (1,0,1,3,7)\,.$$

   Then

$$\mathbf{a}\#\mathbf{b} = (3,6,1,3,5,1,0,1,3,7)\,.$$

2. Let $\mathbf{p} = (p_j)_{j=1}^{16}$ be the list of the first 16 U.S. presidents, in chronological order. Let $\mathbf{q} = (q_j)_{j=1}^{10}$ be the list in chronological order of the first 10 presidents after the 16th one, that is, the list defined by

$$q_j = \text{the } (16+j)\text{−th U.S. president for } j \in \mathbb{N}_{10}\,.$$

   (So, for example, q_1 =Andrew Johnson, q_2 =Ulysses Grant, and so on.)

   Then $\mathbf{p}\#\mathbf{q}$ is the list of the first 26 U.S. presidents, in chronological order. □

And here is the precise definition:

Definition 37. *Let* $\mathbf{a} = (a_j)_{j=1}^{m}$ *and* $\mathbf{b} = (b_j)_{j=1}^{n}$ *be two finite lists. The <u>concatenation</u> of* $\mathbf{a} = (a_j)_{j=1}^{m}$ *and* $\mathbf{b} = (b_j)_{j=1}^{n}$ *is the finite list* $\mathbf{a}\#\mathbf{b}$ *given by*

$$\mathbf{a}\#\mathbf{b} = (c_j)_{j=1}^{m+n}\,, \text{ where } c_j = \begin{cases} a_j & \text{if} \quad j \in \mathbb{N}_m \\ b_{j-m} & \text{if} \quad j \in \mathbb{N} \wedge m+1 \le j \le m+n \end{cases}\,.$$

And here are some of the obvious theorems I announced.

Theorem 80. *If* $\mathbf{a}$ *and* $\mathbf{b}$ *are finite lists of real numbers. Then:*

$$\sum(\mathbf{a}\#\mathbf{b}) = (\sum \mathbf{a}) + (\sum \mathbf{b}), \tag{19.444}$$

$$\prod(\mathbf{a}\#\mathbf{b}) = (\prod \mathbf{a}) \times (\prod \mathbf{b}), \tag{19.445}$$

$$\operatorname{Max}(\mathbf{a}\#\mathbf{b}) = \max\left(\operatorname{Max}\mathbf{a}, \operatorname{Max}\mathbf{b}\right), \tag{19.446}$$

$$\operatorname{Min}(\mathbf{a}\#\mathbf{b}) = \min\left(\operatorname{Min}\mathbf{a}, \operatorname{Min}\mathbf{b}\right). \tag{19.447}$$

Proof. **YOU PROVE THIS.**

Problem 96. ***Prove*** Theorem 80. □

Theorem 81. *Let* $\mathbf{a} = (a_j)_{j=1}^n$, $\mathbf{b} = (b_j)_{j=1}^n$, *be finite lists of real numbers of the same length. Then,*

1. *If*
$$(\forall j \in \mathbb{N}_n) a_j \le b_j$$
then
$$\begin{aligned} \sum \mathbf{a} &\le \sum \mathbf{b} \\ \operatorname{Max}\mathbf{a} &\le \operatorname{Max}\mathbf{b} \\ \operatorname{Min}\mathbf{a} &\le \operatorname{Min}\mathbf{b}. \end{aligned}$$
2. *If all the* a_j *and all the* b_j *are integers, and*
$$(\forall j \in \mathbb{N}_n) a_j | b_j$$
then
$$\prod \mathbf{a} \Big| \prod \mathbf{b}.$$

Proof. **YOU PROVE THIS.**

Problem 97. ***Prove*** Theorem 81. □

Theorem 82. *Let* $\mathbf{a} = (a_j)_{j=1}^n$ *be a finite list of real numbers. Then*

1. $\operatorname{Min}\mathbf{a} \le a_j \le \operatorname{Max}\mathbf{a}$ *for every* $j \in \mathbb{N}_n$.
2. *There exist indices* j_-, j_+ *in* $\mathbb{N}_n$, *such that* $\operatorname{Min}\mathbf{a} = a_{j_-}$ *and* $\operatorname{Max}\mathbf{a} = a_{j_+}$.

Proof. **YOU PROVE THIS.**

Problem 98. ***Prove*** Theorem 81. □

19.3 Prime factorizations

Definition 38. *A prime factorization of a natural number n is a finite list $\mathbf{p} = (p_j)_{j=1}^m$ such that*

(1) p_j is a prime number for every $j \in \mathbb{N}_m$. (That is, all the entries in the list are prime numbers.)

(2) $\prod_{j=1}^m p_j = n$. □

Example 93. The list $(2, 2, 3)$ is a prime factorization of the number 12, because each of the three entries (2, 2, and 3) is a prime number, and the product $2 \times 2 \times 3$ is equal to 12. □

Example 94. The list $(3, 2, 2)$ is also a prime factorization of 12, and is different from the prime factorization $(2, 2, 3)$ of Example 93. □

So the number 12 has at least two different prime factorizations. And yet we want the prime factorization of a natural number to be unique!

To solve this problem we have to introduce the concept of an "ordered prime factorization".

Definition 39. *A finite list $\mathbf{p} = (p_j)_{j=1}^m$ whose entries are real numbers is ordered if*

(ORD) $p_j \leq p_{j+1}$ *for every* $j \in \mathbb{N}_{m-1}$. □

Definition 40. *An ordered prime factorization of a natural number n is a prime factorization $\mathbf{p} = (p_j)_{j=1}^m$ of n which is an ordered list.* □

Example 95. The list $(2, 2, 3)$ is an ordered prime factorization of 12, but the list $(3, 2, 2)$ is not. □

19.4 A correct (and nearly perfect) statement of the FTA

Here, finally, is a correct, nearly perfect[101] statement of the FTA:

Theorem 83. *(A nearly perfect version of the fundamental theorem of arithmetic.) Every natural number n such that $n \geq 2$ has a unique ordered prime factorization.*

[101] I say "nearly perfect" because the statement can be made even nicer and more elegant, thus obtaining a truly "perfect" statement. We will do this later.

19.5 The proof

We have to prove existence and uniqueness of the ordered prime factorization.

The *existence* of a prime factorization of any natural number n such that $n \geq 2$ has been proved before, in Theorem 65 on page 320,

But here we need to prove the existence of an *ordered* prime factorization. Intuitively, this is obvious, because we can take any prime factorization and rearrange the entries putting them in increasing order. More precisely: Let $n \in \mathbb{N}$ be such that $n \geq 2$. Take a prime factorization $\mathbf{p} = (p_j)_{j=1}^m$ of n. (We know that such a factorization exists. Then Rule $\exists_{use}$ enables us to pick one such factorization and call it $\mathbf{p}$.) Then reorder $\mathbf{p}$, by forming a new list $\mathbf{q} = (q_j)_{j=1}^m$ that has the same entries as $\mathbf{p}$, but in increasing order. This gives us an ordered prime factorization of n, proving that such a factorization exists. ***This is not a completely rigorous proof, but the conclusion is fairly obvious, so I will omit the proof at this point. But if you really care about this, and are not satisfied with a nonrigorous proof[102], you can find the proof in the Appendix, on page 442.***

So the existence part of the FTA has been proved.

The uniqueness proof. This is the most delicate part. We have to prove that if we have two ordered prime factorizations $\mathbf{p}$, $\mathbf{q}$, of a natural number n, it follows that $\mathbf{p} = \mathbf{q}$. In other words: we have to assume that

($\Diamond$) We have two finite lists

$$\mathbf{p} = (p_j)_{j=1}^k\,, \qquad \mathbf{q} = (q_j)_{j=1}^\ell\,,$$

such that

(1) all the p_j and all the q_j are prime numbers,

(2) $\mathbf{p}$ and $\mathbf{q}$ are ordered lists (that is, $p_j \leq p_{j+1}$ whenever $j \in \mathbb{N}_{k-1}$, and $q_j \leq q_{j+1}$ whenever $j \in \mathbb{N}_{\ell-1}$),

(3) $\prod_{j=1}^k p_j = \prod_{j=1}^\ell q_j$,

[102] If you take this issue seriously, and want to see a real proof, then I congratulate you: you are thinking like a true mathematician! A true matematician understands that nothing can be justified by saying "it is obvious". If it seems obvious, then either (a) it can be proved easily, or (b) maybe it is not so obvious; maybe it is not even true! Every time something seems obvious to you, you should ask yourself "how can I prove it?". And if you do not know how to prove it, then you should not say it is obvious.

and we want to conclude that

($\Diamond\Diamond$) $\mathbf{p} = \mathbf{q}$.

That is, we want to prove, assuming ($\Diamond$), that

$$k = \ell \wedge (p_j = q_j \text{ for } j = 1, 2, \ldots, k)\,. \tag{19.448}$$

So from now one we assume ($\Diamond$).

First, let prove that $p_1 = q_1$. To prove this, we observe that, since

$$p_1 p_2 \cdots p_k = q_1 q_2 \cdots q_\ell\,,$$

the prime number p_1 divides the product $q_1 q_2 \cdots q_\ell$. Hence, by the generalized Euclid lemma, p_1 divides one of the factors of this product, so we may pick j such that $p_1 | q_j$. Then p_1 is a factor of q_j, so $p_1 = 1$ or $p_1 = q_j$. But p_1 is prime, so[103] $p_1 \neq 1$, So $p_1 = q_j$. But $q_1 \leq q_j$, so $q_1 \leq p_1$.

Similarly, q_1 must equal one of the p_j, and this p_j is $\geq p_1$, so $q_1 \geq p_1$.

Since $q_1 \leq p_1$ and $q_1 \geq p_1$, it follows that $\boxed{p_1 = q_1}$.

We then have, since $p_1 = q_1$,

$$\begin{aligned} p_1 p_2 \cdots p_k &= q_1 q_2 \cdots q_\ell \\ &= p_1 q_2 \cdots q_\ell\,, \end{aligned}$$

so $p_1 p_2 \cdots p_k = p_1 q_2 \cdots q_\ell$, from which it follows that

$$p_2 \cdots p_k = q_2 \cdots q_\ell\,.$$

So we find ourselves in the same situation we started with, except that now we have p_2, q_2 in the role previously played by p_1, q_1. So, repeating the same argument, we get $p_2 = q_2$ and then we can go on and repeat the argument once more and prove that $p_3 = q_3$, and so on.

[103] ***Notice how important it is that in the definition of "prime number" (definition 4, on page 20) we included the requirement that, for p to be prime, p has to be > 1.*** This is the step where that condition is used. As explained in section 2.3.1, or page 21, if we had decided to count 1 as a prime number, then the Fundamental Theorem of Arithmetic would not be true. What would fail is the uniqueness part. For example, we could take $k = 2$, $\ell = 3$, $p_1 = 2$, $p_2 = 3$, $q_1 = 1$, $q_2 = 2$, and $q_3 = 3$, and we would get $p_1 p_2 = q_1 q_2 q_3$, with p_1, p_2, q_1, q_2, q_3 prime, $p_1 \leq p_2$, and $q_1 \leq q_2 \leq q_3$, but it is not true that $\ell = k$ and $p_1 = q_1$ and $p_2 = q_2$. So it is not surprising that, since the condition "$p \neq 1$" is needed for the uniqueness part of the FTA to be valid, it is precisely in the proof of the uniqueness part of the FTA that this condition is used. And the step where it is used is preciely here.

However, we know that "and so on" is problematic, and the rigorous way to do an "and so on"' argument is with a proof by induction. So let us do a proof by induction.

What we have done so far is show that we can prove that $p_1 = q_1$, and then go from that to $p_2 = q_2$, then go from that to $p_3 = q_3$. So this suggests that, for our induction, we could use the predicate $P(n)$, where $P(n)$ stands for "$p_1 = q_1 \wedge p_2 = q_2 \wedge \cdots \wedge p_n = q_n$".

There is, however, a minor problem with this idea:

- $P(n)$ only makes sense for n if $p_1, p_2, \ldots, p_n$ and $q_1, q_2, \ldots, q_n$ are defined, that is, if $n \le k$ and $n \le \ell$.
- But to do induction we need a predicate that makes sense for every $n \in \mathbb{N}$.

So we modify the previous $P(n)$ a little bit and use instead the following choice for $P(n)$:

(*) We let $P(n)$ be the predicate

$$\text{if } \ n \le k \ \text{ and } \ n \le \ell \ \text{ then } \ p_j = q_j \ \text{ for } \ j = 1, 2, \ldots, n\,. \tag{19.449}$$

That is,

$$P(n) \ \text{ stands for: } \quad (n \le k \wedge n \le \ell) \Longrightarrow (\forall j \in \mathbb{N}_n) p_j = q_j\,. \tag{19.450}$$

(The virtue of this precdicate is that when $n > k$ or $n > \ell$, the premise "$n \le k \wedge n \le \ell$)" is false, so $P(n)$ is true. and we don't need to worry about the issue whether p_n or q_n is well defined.)

Let prove $(\forall n \in \mathbb{N}) P(n)$ by induction.

Basis step. We want to prove $P(1)$, that is,

$$(1 \le k \wedge 1 \le \ell) \Longrightarrow p_1 = q_1\,. \tag{19.451}$$

But we have already proved that $p_1 = q_1$. So (19.451) is true, and we have proved $\boxed{P(1)}$.

Inductive step. We want to prove that

$$(\forall n \in \mathbb{N}) \Big(P(n) \Longrightarrow P(n+1) \Big)\,. \tag{19.452}$$

Let $n \in \mathbb{N}$ be arbitrary. Assume $P(n)$.

We want to prove $P(n+1)$. That is, we want to prove

$$(n+1 \le k \wedge n+1 \le \ell) \Longrightarrow \big(p_1 = q_1 \wedge \cdots \wedge p_{n+1} = q_{n+1}\big)\,. \quad (19.453)$$

To prove the implication (19.453) we assume the premise and try to prove the conclusion.

Assume that $\boxed{n+1 \le k \wedge n+1 \le \ell}$.

Then $n < k$ and $n < \ell$, so in particular $n \le k \wedge n \le \ell$.

Since we are assuming $P(n)$, we know that

$$(n \le k \wedge n \le \ell) \Longrightarrow \big(p_1 = q_1 \wedge \cdots \wedge p_n = q_n\big)\,. \quad (19.454)$$

But we know that $n \le k \wedge n \le \ell$, which is the premise of the implication (19.454).

Then Rule $\Longrightarrow_{use}$ (the Modus Ponens rule) allows us to go to the conclusion of (19.454), i.e.,

$$p_1 = q_1 \wedge \cdots \wedge p_n = q_n\,. \quad (19.455)$$

Since $n < k$ and $n < \ell$, the equality $p_1 p_2 \cdots p_k = q_1 q_2 \cdots q_\ell$ can be rewritten as

$$p_1 p_2 \cdots p_n p_{n+1} \cdots p_k = q_1 q_2 \cdots q_n q_{n+1} \cdots q_k q_{k+1} \cdots q_\ell\,,$$

and, since $p_j = 1$ for $j = 1, \ldots, n$, this says

$$p_1 p_2 \cdots p_n p_{n+1} \cdots p_k = p_1 p_2 \cdots p_n q_{n+1} \cdots q_k q_{k+1} \cdots q_\ell\,,$$

from which it follows that

$$p_{n+1} \cdots p_k = q_{n+1} \cdots q_k q_{k+1} \cdots q_\ell\,. \quad (19.456)$$

We then repeat the same argument used earlier to prove that $p_1 = q_1$ and conclude that $p_{n+1} = q_{n+1}$. (The prime p_{n+1} divides the product $q_{n+1} \cdots q_k q_{k+1} \cdots q_\ell$, so it is equal to one of the factors; but this factor is $\ge q_{n+1}$, so $p_{n+1} \ge q_{n+1}$; similarly, $q_{n+1} \ge p_{n+1}$; and then $p_{n+1} = q_{n+1}$.)

Since we already know that $p_j = q_j$ for $j = 1, \ldots, n$, we have proved that

$$p_1 = q_1 \wedge \cdots \wedge p_{n+1} = q_{n+1}\,, \quad (19.457)$$

Since we have proved (19.457) assuming that $n+1 \leq k \wedge n+1 \leq \ell$, we have proved that

$$(n+1 \leq k \wedge n+1 \leq \ell) \Longrightarrow \big(p_1 = q_1 \wedge \cdots \wedge p_{n+1} = q_{n+1})\big)\,. \quad (19.458)$$

That is, we have proved $P(n+1)$.

Since we have proved $P(n+1)$ assuming $P(n)$, we have proved the implication $P(n) \Longrightarrow P(n+1)$.

Since we have proved $P(n) \Longrightarrow P(n+1)$ for arbitrary $n \in \mathbb{N}$, it follows that $\boxed{(\forall n \in \mathbb{N})\Big(P(n) \Longrightarrow P(n+1)\Big)}$.

This completes the inductive step. Since we have also proved $P(1)$, we can conclude, thanks to the PMI, that $(\forall n \in \mathbb{N})P(n)$.

End of the uniqueness proof. Now that we have proved that $P(n)$ is true for every $n \in \mathbb{N}$, we can conclude our uniqueness proof.

Let $\nu = \min(k, \ell)$, so ν is the smallest of k and ℓ.

Then $\nu \in \mathbb{N}$, so $P(\nu)$ is true.

But $P(\nu)$ says

$$\text{if } \ \nu \leq k \ \text{ and } \ \nu \leq \ell \ \text{ then } \ p_j = q_j \ \text{ for } \ j = 1, 2, \ldots, \nu\,. \quad (19.459)$$

But $\nu \leq k$ and $\nu \leq \ell$, so we can conclude that

$$p_j = q_j \ \text{ for } \ j = 1, 2, \ldots, \nu\,. \quad (19.460)$$

We are now going to prove that $\ell = k$. Suppose $\ell > k$. Then $\nu = k$, and the formula $p_1 p_2 \cdots p_k = q_q q_2 \cdots q_\ell$ can be rewritten as

$$\begin{aligned} p_1 p_2 \cdots p_k &= q_1 q_2 \cdots q_k q_{k+1} q_{k+2} \cdots q_\ell \\ &= p_1 p_2 \cdots p_k q_{k+1} q_{k+2} \cdots q_\ell\,. \end{aligned}$$

Hence

$$q_{k+1} q_{k+2} \cdots q_\ell = 1\,.$$

but this is impossible, because the product $q_{k+1} \cdots q_\ell$ is a product of at least one prime[104], so the product is > 1.

Hence it is not true that $\ell > k$. A similar argument shows that it cannot happen that $\ell < k$. So $\boxed{\ell = k}$.

Since $\ell = k$, ν equals k as well, and then formula (19.460) tells us that $\boxed{p_j = q_j \text{ for } j = 1, 2, \ldots, k}$.

Ths completes the proof. **Q.E.D.**

[104] There is at least one prime in this product because $\ell > k$.

19.5.1 The perfect statement of the FTA

Mathematicians like to have their theorems as simple and general as possible. The FTA, as we have stated it, has a condition that makes it inelegant, namely, the requirement that $n \geq$.

Wouldn't it be nicer if we could just say

> **Theorem 84** *(The fundamental theorem of arithmetic.) Every natural number has a unique ordered prime factorization.*

?

This would clearly be more elegant, wouldn't it? It's much simpler than our previous version, and it is also more general, because it applies to all natural numbers, even to the number 1.

But, of course, just because a statement is nice, it doesn't mean that it is true.

Is our new statement of the FTA true? The answer is "yes", but we have to be careful about what this means.

Notice that the only difference between the previous statement of the FTA and our new statement is that the new statement says that the number 1 also has a unique ordered prime factorization. And we have to ask the obvious question: *what is that factorization?*

The answer is: *the ordered prime factorization of* 1 *is the empty list.* Let me explain.

First of all, until now we said that every list has a length, and that this length is a natural number. We now change that, and add a new list: ***the empty list***.

The empty list is a list of length zero, that has no entries whatsoever. We use the symbol $\emptyset$ to denote this list[105].

And we can also think of the empty list as the list $(a_j)_{j=1}^{0}$, because there are no values of j such that $1 \leq j$ and $j \leq 0$, so the list $(a_j)_{j=1}^{0}$ has no entries.

Then the following is true:

Proposition 3. *The empty list is an ordered list of primes.*

[105] You may worry that "$\emptyset$" already stands for the empty set. You need not worry. If one does things carefully, it turns out that the empty set and the empty list truly are the same thing, so it is perfectly all right to use "$\emptyset$" both to denote the empty set and to denote the empty list. But it takes some work to establish this, so for the moment just accept that the empty list is called "$\emptyset$".

This can be rigoruously proved as follows.

Proof. First, we want to prove that $\emptyset$ is a list of primes.

Write the empty liist $\emptyset$ as $(p_j)_{j=1}^0$.

We have to prove that

$$(\forall j)(j \in \mathbb{N}_0 \Longrightarrow p_j \text{ is a prime number}) \tag{19.461}$$

where "p_j" stands for "the j-entry of the empty list".

So let j be arbitrary. We want to prove that

$$j \in \mathbb{N}_0 \Longrightarrow p_j \text{ is a prime number}\,. \tag{19.462}$$

But $\mathbb{N}_0$ is the empty set, so $\mathbb{N}_0$ has no members, and then "$j \in \mathbb{N}_0$" is false, no matter who j might be.

Since "$j \in \mathbb{N}_0$" is false, the implication (19.462) is true.

So we have proved (19.462), for arbitrary j, And then we have proved (19.461).

We can use a similar argument to prove that $\emptyset$ is an ordered list. (Sketch of the argument: we have to prove that "if $j \in \mathbb{N}_0$ and $j+1 \in \mathbb{N}_0$ then $p_j \leq p_{j+1}$". And this is true because it is an implication with a false premise.) **Q**.**E**.**D**.

Finally, it turns out that $\boxed{\prod_{j=1}^0 p_j = 1}$. If you have trouble believing this, I will give you three reasons:

Reason No.1: $\prod_{j=1}^0 p_j = 1$ because in these notes we defined $\prod_{j=1}^0 p_j$ to be equal to 1, when we gace the inductive definition of "$\prod$".

Reason No.2: $\prod_{j=1}^0 p_j = 1$ because mathematicians have agreed that this is so. In other words, the statement "$\prod_{j=1}^0 p_j = 1$" is ***true by convention***, because mathematicians have agreed that the product of the empty list is equal to one[106].

Reason No.3: Mathematicians are reasonable people, so if we decided that $\prod_{j=1}^0 p_j = 1$ we must have had a good reason.

[106]This is like many other conventions. Why is Pluto not a planet? Because astronomers have decided that it is isn't. Why is 1 not a prime number? Because mathematicians have decided that it isn't. Why do we drive on the right side of the street? Because at some point it was decided (in the U.S and many other countries, but not in all countries) that the right side of the street is the side on which people should drive. Why are cows called "cows" rather than, say, "zebras", or "tables"? Because English-speaking people have agreed that that is the name of those animals.

Here is the reason. The inductive definition of "$\prod$" tells us that

$$\prod_{j=1}^{n+1} p_j = \Big(\prod_{j=1}^{n} p_j\Big) p_{n+1} \tag{19.463}$$

if n is a natural number. This means that

$$\prod_{j=1}^{n} p_j = \frac{\prod_{j=1}^{n+1} p_j}{p_{n+1}} \tag{19.464}$$

for $n \in \mathbb{N}$. Now suppose we want to make Formula (19.464) also true for $n = 0$. Then we must have

$$\prod_{j=1}^{0} p_j = \frac{\prod_{j=1}^{1} p_j}{p_1} \,. \tag{19.465}$$

But

$$\prod_{j=1}^{1} p_j = p_1 \,.$$

So we must have

$$\prod_{j=1}^{0} p_j = \frac{p_1}{p_1} = 1 \,. \tag{19.466}$$

This is not a rigorous proof. But it is an argument showing that the convention that $\prod_{j=1}^{0} p_j = 1$ is a reasonable one.

In any case, ***once you agree that*** $\prod_{j=1}^{0} p_j = 1$ ***follows that our nicer version of the FTA is true.***

20 Definitions: how you should write them and how you should not write them

20.1 An example of a correctly written definition

Suppose you don't know what a prime number is. And suppose you are asked whether the numbers 1, 2, 6, 7, 10, 12, are "prime". Then you will probably not be able to answer the question, because you don't know what a "prime number" is. So you would answer with a question: *what is a prime number"*, or *what does it mean for a number to be prime?*

To answer such a question, you need to know the **definition** of "prime number".

And here is the definition:

DEFINITION OF "PRIME NUMBER"

Let n be a natural number. We say that n is prime if $n \neq 1$ and the only natural numbers that are factors of n are 1 and n.

And here is another, equally correct, definition of "prime number":

DEFINITION OF "PRIME NUMBER", VERSION II

A natural number n is prime if $n \neq 1$ and every natural number that is a factor of n is either equal to 1 or to n.

And here is a third, also completely correct, definition of "prime number":

DEFINITION OF "PRIME NUMBER", VERSION III

A natural number n is prime if $n \neq 1$ and
$(\forall q \in \mathbb{N})\Big(q|n \Longrightarrow (q = 1 \vee q = n)\Big)$.

And, finally, here is a fourth completely correct definition of "prime number":

DEFINITION OF "PRIME NUMBER", VERSION IV:

An integer n is a prime number if $n > 1$ and $(\forall q \in \mathbb{N})\Big(q|n \implies (q = 1 \vee q = n)\Big)$.

20.2 How not to write a definition

Let us look now at some bad ways of writing the definition "prime number".

The examples I am going to give you are representative of things students often write in exams. ***You should read these examples carefully, and then read the explanation of why these definitions are bad, so that you will learn not to write that way.***

Some of the definitions below are truly horrendous (and would get zero points on a scale from 0 to 10), while others are not 100% wrong but are not entirely correct either, and may get 5 points on a 0-10 scale, or maybe in some cases even 6 or 7. But ***you should understand why those definitions are bad, so you can learn how to write definitions correctly and get*** 10 ***points out of*** 10***.***

Bad Definition 1. Prime number is when you cannot divide by any number other than by the number itself. □

Bad Definition 2. A prime number is a number that cannot be divided by any number other than 1 and itself. □

Bad Definition 3. A prime number is a natural number that cannot be divided by any number other than 1 and itself. □

Bad Definition 4. A prime number is a natural number such that the only factors of the number are 1 and the number itself. □

Bad Definition 5. A prime number is a natural number such that the only factors of n are 1 and n. □

Bad Definition 6. A prime number is a natural number such that the only natural numbers that are factors of n are 1 and n. □

Bad Definition 7. A prime number is a natural number such that $n > 1$ and the only natural numbers that are factors of n are 1 and n. □

Bad Definition 8. A prime number is a natural number n such that $n > 1$ and the only natural numbers that are factors of n are 1 and n. □

20.2.1 Analysis of bad definitions

Let us analyze our eight "bad definitions" and explain why they are bad.

The main question that we will ask, and the question that you should always ask, is: ***using this definition, can I tell correctly if an object is what the definition says it is supposed to be?*** (In this case, ***can I tell correctly if an object is a prime number or not?***)

Notice that this question really amounts to two questions:

(I) ***Can I tell?***, that is, *does the definition tell me precisely what to do in order to find out if the answer is "yes" or "no"?*

(II) Can I tell ***correctly***?, that is, *when I do what the definition tells me to do, do I get the right answers?*

Question (I) is the ***precision and clarity*** question: does the definition tell me cearly and precisely what I am supposed to do in order to find out the answer?

Question (II) is the ***correctenss*** question: If I do what the definition tells me to do, do I get the right answer?

These two questions are different. For example, if I were to define "prime number" as follows:

Bad Definition 9. A prime number is a natural number that is divisible by 2. □

Then this definition is completely clear and precise. It tells me that in order to find out if a number is prime, I have to see if it is divisible by 2. The problem with this definition is that it does not satisfy the ***correctness*** condition: if I apply the definition, say, to the number 6, I find that 6 is divisible by 2, so according to this definition 6 is prime, which is not true.

To assess a definition, you should always ask these two questions: is the definition clear and precise, so that when I want to apply it I know exactly what to do? And is it correct, in the sense that it gives me the right answers?

And, in order to answer the correctness question, you should ***test*** your definition by applying it to several examples and seeng whether it gives the right answer.

The simplest and most convincing way to establish that a definition is wrong is to give an example of something for which the definition gives thw rong answer. This is what we did when we disucssed the

You should always ask these two questions, ***especially about definitions you have written yourself.*** And if what you wrote does not meet the two requirements of (1) precision and clarity and (2) correctness, then your definition is not acceptable and you must work on it until you get it right.

Now let us look at the eight bad definitions in our list.

1. Bad Definition 1 says: *Prime number is when you cannot divide by any number other than by the number itself.*

   This is truly atrocious. Let us see why.

   - First of all, when you say "prime number is", you are suggesting that "prime number" is a condition of the world, such as "chaos", or "peace". You can say something like "peace is when people are not fighting", or "chaos is when there is utter confusion". Even these sentences are very bad English, but you can more or less figure out what they mean. (For example, when you see that people are fighting, you would say that "there is no peace here", and when people stop fighting, you would say "now there is peace".) Much better ways to say these things would be: "Peace is the absence of war or other hostilities", or "Peace is a state of affairs in which people are not fighting".
   - But in the case of "prime number", the "prime number is when" construction does not make sense. Being a prime number is not some kind of state of affairs. It is a property of a specific kind of object, namely, numbers. So one has to use much more precise language, and start the definition with "A prime number is", or "A number is prime if".

If a definition starts with "such and such is when..." you can be sure it is wrong:

- *"Prime number is when..." is wrong.*
- *"Divisible is when..." is wrong.*
- *"Even number is when..." is wrong.*
- *"Power set is when..." is wrong.*
- *"Subset is when..." is wrong.*
- *"Intersection is when..." is wrong.*

A correct definition of "prime number" should start in one of the following ways:

- "A prime number is a natural number n such that"
- "Let n be a natural number. Then n is a prime number if"
- "Let n be a natural number. We say that n is prime if"
- "A natural number n is prime if"

In other words: ***at the beginning of the definition you have to introduce the object or objects that you will be talking about***. In this example, you do this by indicating that you will be talking about a natural number, not about a real number or a cow or a fish or a river. And you may give that natural number a name, such as n.

2. Bad Definitions 1 and 2 talk about "numbers". We have already quoted Bad Definition 1, and Bad definition 2 says: *A pime number is a number that cannot be divided by any number other than* 1 *and itself.*

   This definition does not pass the "can I tell?" test. It tells me that to be a prime number an object has to be a "number".

   But ***"number" is a vague concept***, because there are lots of different kinds of "numbers", so when you say "number" you could mean

"natural number" (that is, the kind of number that you are used to calling "whole number"), "integer", "rational number", "real number", "complex number", or lots of other kinds of numbers that exist.

Never say "number" unless it is clear what kind of "number" you are talking about.

If I want to follow Bad Definition 2, the, when I am given a thing and want to find out if that thing is a prime number, the first I thing I have to do is find out if it is a "number". But I cannot do that because I don't know what a "number" is. So the definition fails the "can I tell?" test.

In a correct, intelligible definition, when you talk about a 'number", you have to make it clear what you mean by "'number".

This can be made clear in at least three ways:

- You can just say what kind of number your number is supposed to be. (For example, you could say "let n be a natural number", or "let n be an integer", or "let n be a rational number", or "let n be a real number".)
- You can make it clear at the beginning of your text that the word "number" is always going to mean "integer", or "real number", or whatever. If you do so, then you don't need to repeat that you mean "integer", or "real number", or whatever, every time you say "number".
- You may want to talk about different kinds of numbers simultaneously. And, in order to do that, you may declare, at the beginning of your text, that, for example, "in this chapter, the letters m, n, p, q will always stand for natural numbers, and the letters x, y, z, u, v, w will stand for real numbers".

3. Bad definitions 1, 2, and 3, talk about "dividing by numbers", and tell me that a number is prime if it cannot be divided by certain numbers. But this is very confusing.

   - Actually, ***any number can be divided by any number (except zero)***. For example, I can divide 7 by 5, getting as a result the number $\frac{7}{5}$.

- So the issue is not whether "we can divide", because wwe can almost always do that, but ***what kind of result we get when we divide***.
- When Bad Definition 3 tells me that I should see if a number "can be divided by numbers other that 1 and the number itself", then I could try to apply the definition, for example, to the number 3, and I would immediately see that 3 can be divided by lots of numbers other than 1 and 3: I can divide 3 by 2 (and the result is $\frac{3}{2}$), I can divide 3 by 7 (and the result is $\frac{3}{7}$), I can divide 3 by 29 (and the result is $\frac{3}{29}$), and so on.

4. Bad Definitions 4 and 5 are a little bit better. Rather than talk about "dividing", they talk about "factors", which is more precise. because we have a precise definition of "factor".

   But that is not good enough. According to the definition of "factor", a factor of an integer a is an integer b such that there exists an integer k for which $a = bk$. So, when Bad Definition 5 says that

   > *A prime number is a natural number such that the only factors of n are 1 and n.*

   then ***this definition fails the correctness test***: ***according to this definition*** 2 ***is not prime***, becauuse 2 as other factors in addition to 1 and 2. Indeed, -1 and -2 are factors of 2 as well, since $2 = (-1) \times (-2)$ and $2 = (-2) \times (-12)$.

5. Bad Definition 6 is much better. It says that

   > *A prime number is a natural number such that the only natural numbers that are factors of n are 1 and n.*

   This is quite close, but ***this definition still fails the correctness test, because it gives us wrong answers***. Indeed, according to this definition 1 is prime. But this is wrong: 1 is not prime[107].

6. With Bad Definition 7 we enter, for the first time, the "partial credit" zone. This definition is essentially correct, but it is not well written. It says that

[107] Why is 1 not prime? For the same reason why Pluto is not a planet. Mathematicians have decided not to call 1 "prime", exactly as astronomers have decided not to call Pluto a planet. But this decision was made for good reasons, that will be discussed later in this course.

> *A prime number is a natural number such that $n > 1$ and the only natural numbers that are factors of n are 1 and n.*

The problem with this is that the defintion talks about "n" but does not tell us who this "n" is. ***In a mathematical text, when you refer to an object using a letter name, this name has to be introduced first.***

7. Bad Definition 8 does this: the symbol "n" is properly introduced when we are told that

   > *A prime number is a natural number n such that $n > 1$ and the only natural numbers that are factors of n are 1 and n.*

8. So Bad Ddefinition 8 is nearly perfect. What is missing? Only one thing: ***in a definition, the word or phrase being defined must be highlighted is some way, to indicate that we are defining that word or phrase.*** And when we write by hand the way we highlight is by underlining. So, for example, in a definition of "'prime number" the words "prime numebr" have to be underlined. And if we do that we get a correct definition. *A <u>prime number</u> is a natural number n such that $n > 1$ and the only natural numbers that are factors of n are 1 and n.*

20.2.2 Always highlight the definiendum

When you write a definition, you are defining a particular word or phrase. That word or phrase is called the *definiendum*. (This just means "the thing being defined.") ***The definiendum should always be highlighted.***

In books, the authors do this by using Italics, or Boldface. But when we write by hand, it is hard to do Italics or Boldface, so we use underlining.

Look, for example, at any definitions you want in our textboook. Just open the book at random, at any page, and look at the definitions on that page. And, for each definition, ask yourself "what is this definition the definition of?" And, invariably, you will see that the term or phrase being defined is in **boldface**. (This is not just a peculiarity of our textbook. It's done in every Mathematics book.) In my lecture notes, I use underlining rather than boldface. And when you write your homework or your exams, or when I write on the blackboard, it's hard to do italics or boldface, so I use underlining instead, and you should do the same.

20.3 The general formats for definitions

In a definition, the word, symbol or phrase whose meanign we are trying to define is called the definiendum.

20.3.1 Step 1: Find out if the definiendum is a term or a sentence, and what its arguments are

In order to know how to write a definition of something, we first have to figure out two things:

1. Whether the definiendum is a ***term*** or a ***sentence***.
2. What the ***arguments*** of the definiendum are.

Recall that

- A ***term*** is a word or symbol or phrase that stands for a thing. Terms are essentially the same things that in your English or linguistics classes you may have called "noun phrases".
- A ***sentence*** is a word or symbol or phrase that makes an assertion that can be true or false. Sentences are essentially the same things as "predicates", or "statements".
- Terms and sentences have ***values***.
- The value of a term is the thing the term stands for. For example the term "New York City" is New York City.
- The value of a sentence is its truth value. For example, the sentence "New York City is the capital of New York State" has the truth value "false", because it is not true, but the sentence "Albany is the capital of New York State" has the truth value "true", because it is true.
- If a term or sentence contains variables, then the term or sentence only has a value, or truth value, is the variables that occur in it have been assigned values. For example,
  - the term "$x + y$" contains two variables, x and y. If we assign values to these variables, by saying something like "let $x = 5$, $y = 3$", then the term "$x + y$" has the value 8.
  - the sentence '$x + y = z$" contains three variables, x, y, and z. If we assign values to these variables, by saying something like "let $x = 5$, $y = 3$, $z = 4$", then the sentence "$x + y = z$" has the truth value "false", because $5 + 3$ is not equal to 4.

20.4 Step 2: Introduce the arguments

You must start your definition by introducing the arguments.

For example:

- If you want to define "prime number". then you will see, first of all, that the definiendum is a sentence, "something is a prime number". And it has one argument, because we say things such as "n is a prime number". What you want to explain to the readers is how to tell what the truth value of the definiendum is for any given value or values of the arguments. That is, you want to tell the readers under what conditions they should call a number n "prime", that is, when they should say "n is prime". So you definition must start by saying something like "Let n be an integer", or "let n be a natural number", or "let n be a real number". (Eventually, n will turn out to be a natural number anyhow. So you could start your definition by requiring n to be a natural number. But you can also require n to be an integer, and let the second part of the definition force n to be a natural number, for example by putting the requirement that $n > 1$. And you could even start by requiring n to be a real number, and then say later: "we say that n is a a prime number if it is a natural number such that ...".)

- "Divisible" is a two-argument sentence, because we say things such as "m is divisible by n", and these things are true or false. So in the definition of "divisible" you want to tell the readers under what conditions they should say of two numbers m, n that "m is divisible by n". And you must start by introducing the two numbers m and n, by saying something like "Let m, n be integers".

- "Union" is a two-argument term, because we talk about "the union of two sets A, B", and that union is a thing, namely, a set. So in the definition of "union" you want to tell the readers who the set $A \cup B$ is, if we are given two sets A, B. So you must start by introducing the two sets A and B, by saying something like "Let A, B be sets".

- "Subset" is a two-argument sentence, because we say things such as "A is a subset of B", and this sentence is true or false. So in the definition of "subset" you want to tell the readers under what conditions they should say that "$A \subseteq B$" is true, if we are given two sets A, B. And you must start by introducing the two sets A and B, by saying something like "Let A, B be sets".

- "Power set" is a one-argument term, because we talk about "the power set of a set A", and that power set is a thing, namely, a set. So in the definition of "power set" you want to tell the readers who the set $\mathcal{P}(A)$ is, if we are given a set A. So you must start by introducing the set A, by saying something like "Let A be a set".

- "Derivative" is more complicated, because there are two different concepts of derivative:

  - We talk about "the derivative of a function f at a point a." This is a two-argument term: the derivative of f at a is a real number. So your definition of "derivative of a function at a point" must start by saying something like "Let f be a function and let a be a real number".
  - We talk about "the derivative of a function f." This is a one-argument term: the derivative of a function f is another function, usually called f'. So your definition of "derivative of a function" must start by saying something like "Let f be a function".

- "married" is also complicated, like "derivative". because there are two different concepts of "married":

  - We talk about "two people begin married to each other." This is a two-argument sentence: if x and y are people, then "x and y are married to each other" can be true or false. So your definition of "x and y are married to each other" must start by saying something like "Let x, y be two persons".
  - We talk about one person being married, and say things like "x is married." This is a one-argument sentence. So your definition of "married" must start by saying something like "Let x be a person".

20.5 Step 3: Tell the readers how to find the value of the definiendum

Now that you have introduced the arguments, you have to tell your readers how they can determine the value of the definiendum for those arguments. That value will be a thing if the definiendum is a term, and a truth value if the definiendum is a sentence.

For example:

- In the definition of "prime number", after you have said, for example, "Let n be a natural number", you have to tell the readers how to figure out the value of the definiendum, for n. In this case, the definiendum is the sentence "n is prime", so you you have to tell the readers what has to happen that will make that sentence true. You can say, for example: "We say that n is a prime number if $n \neq 1$ and $(\forall m \in \mathbb{N})\Big(m|n \Longrightarrow (m = 1 \vee m = n)\Big)$".

- In the definition of "divisible". after you have said "Let m, n be integers", you have to tell the readers how to figure out the value of the definiendum, for m and n. In this case, the definiendum is the sentence "m is divisible by n", so you have to tell the readers what has to happen that will make them say that the sentence is true. You can say, for example: "We say that m is divisible by n if $(\exists k \in \mathbb{Z})m = nk$.".

- In the definition of "union". after you have said "Let A, B be sets", you have to tell the readers how to figure out the value of the definiendum, for A and B. In this case, the definiendum is the term "$A \cup B$", which is the name of a set. So you you have to tell the readers who that set is, by saying, for example: "the union of A and B is the set $A \cup B$ given by $A \cup B = \{x : x \in A \vee x \in B\}$."

- In the definition of "power set". after you have said "Let A be a set", you have to tell the readers how to figure out the value of the definiendum, for the set A. In this case, the definiendum is the term "$\mathcal{P}(A)$", which is the name of a set. So you you have to tell the readers who that set is, by saying, for example: "the power set of A is the set $\mathcal{P}(A)$ given by $\mathcal{P}(A) = \{X : X \subseteq A\}$."

Problem 99. ***Analyze critically*** (and, in particular, assign a grade on a scale from[108] 0 to 10) to each of the following definitions[109]:

1. *Definition of "divisible"*: Divisible is when it can be divided.
2. *Definition of "divisible"*: A number n is divisible if it can be divided evenly into many parts.
3. *Definition of "divisible"*: A number n is divisible if it can be divided evenly into many parts.

[108] You are allowed to give negative grades like -300 for particularly atrocious definitions. And, since the authors of these definitions are just figments of my imagination, not real students, you don't have to be politically correct and worry about the danger that you might hurt their feelings, and should feel free to be very harsh.

[109] At least two of the fifteen definitions in our list deserve a huge negative grade.

4. *Definition of "divisible"*: Let m, n be integers. We say that m is divisible by n if $\frac{m}{n}$ is an integer.
5. *Definition of "even integer"*: An even integer is $n = 2k$.
6. *Definition of "even integer"*: An even integer is $n = 2k$, where k is an integer.
7. *Definition of "even integer"*: An even integer is $n = 2k$, where k is an integer.
8. *Definition of "even integer"*: An even integer is an integer such that $n = 2k$, where k is an integer.
9. *Definition of "even integer"*: An even integer is an integer such that $n = 2k$ for some integer k.
10. *Definition of "even integer"*: An even integer is an integer n such that $n = 2k$ for some integer k.
11. *Definition of "even integer"*: An even integer is an integer n such that $(\exists k \in \mathbb{Z})n = 2k$.
12. *Definition of "even integer"*: An even integer is an integer n where $n = 2k$ for some integer k.
13. *Definition of "prime number"*: Prime is when it cannot be divided by anything.
14. *Definition of "prime number"*: A prime number is a number that is not divisible by anything.
15. *Definition of "prime number"*: A prime number is a natural number n such that n has exactly two positive integer factors. □

21 Sets

The language of ***sets*** was introduced into mathematics in the 19th century, when the great mathematician ***George Cantor*** (1845-1918) almost single-handedly created ***Set theory***.

You should read the article "A history of set theory", in MacTutor.

Today, set theory is not only an important branch of mathematics, but the foundational pillar on which all of mathematics rests. Most mathematicians no longer ask questions that they used to ask, such as "what is a natural numebr?", or "what is a real number?", or "what is a function?", because they think that all these objects are just special kinds of sets.

This does not mean that they have answered those questions. It just means that they have reduced those questions to just one question: what is a set? Once you know what a set is, then all the other questions are answered.

As for the fundamental question "what is a set?", I am not going to answer it here. What I am going to do is start telling you about sets, until you get used to working with them and talking about them. The question about the ultimate nature of sets will remain unanswered.

21.1 What kind of thing is a set?

Sets are things that we invent in order to combine several objects and form with them a single thing, so that we can talk about the objects as one thing, a "collective entity".

This "grouping" operation, of forming a single thing out of several things, is something we perform very often, using different words, called "collective nouns", to create these collective objects.

Here are some examples.

1. ***Crowds.*** When you see a number of people standing together and shouting something (say, "long live the Queen"), you create a single thing, called "the crowd", so that, instead of saying

   the people are shouting "long live the Queen"

   you can use the collective noun "crowd" and say

   the crowd is shouting "long live the Queen"

*Notice that "the people" have become a single object, "the crowd". So, instead of using the verb in plural ("the people **are** shouting") when you talk about the people, you use the verb in singular ("the crowd **is** shouting") when we talk about the crowd.*

2. ***Flocks of birds.*** When we see a number of birds flying in formation, we create an entity called "the flock", so that, instead of saying

   I see several birds, and they are flying East,

   we can use the collective noun "flock" and say

   I see a flock of birds, and it is flying East.

   *Notice that "the birds" have become a single object, "the flock". So, instead of using the verb in plural ("the birds **are** flying") when we talk about the birds, we use the verb in singular ("the flock **is** flying") when we talk about the flock.*

3. ***Orchestras.*** When several musicians are playing together, we introduce into our discourse the collective noun "orchestra", so that, instead of saying

   The musicians are playing

   we can use the collective noun "orchestra" and say

   The orchestra is playing.

   *Once again, "the musicians" have become a single object, "the band". So, instead of using the verb in plural ("the musicians **are** playing") when we talk about the musicians, we use the verb in singular ("the orchestra **is** playing") when we talk about the orchestra.*

4. ***Juries.*** When several people are brought together to sit in judgemebnt and decide if a defendant is guilty, the people are called ***jurors***, and are said to be members of the ***jury***.

   And we say things like

The jurors ***find*** the defendant guilty

or

The jury ***finds*** the defendant guilty.

Once again, when we talk about "the jurors" we use the verb in plural ("find") but when we talk about "the jury" itself we use the verb in singular ("finds") because the jury is a single object.

5. ***The sets*** $\mathbb{N}$***,*** $\mathbb{Z}$***, and*** $\mathbb{R}$***.*** When numbers of a certain kind are discussed together, we create entities called $\mathbb{N}$ ("the set of all natural numbers"), $\mathbb{Z}$ ("the set of all natural integers"), $\mathbb{R}$ ("the set of all real numbers"), so that, instead of saying

there are infinitely many natural numbers

we can use the collective noun "$\mathbb{N}$" and say

the set $\mathbb{N}$ is infinite.

Similarly, instead of saying

all integers are real numbers,

we can use the collective nouns "$\mathbb{N}$" and "$\mathbb{Z}$" and say

$\mathbb{Z}$ is a subset of $\mathbb{R}$.

And, instead of saying

the real numbers form a complete ordered field,

we can use the collective noun "$\mathbb{R}$" and say

$\mathbb{R}$ is a complete ordered field.

*Notice that "the natural numbers", "the integers", and "the real numbers" have become single objecta, "$\mathbb{N}$". "$\mathbb{Z}$". "$\mathbb{R}$". So, instead of using verbs in plural ("there **are** infinitely many natural numbers", "all integers **are** real numbers", "the real numbers **form**..."), when we talk about the numbers, we use verbs in singular ("the set $\mathbb{N}$ **is** infinite", "$\mathbb{Z}$ **is** a subset of $\mathbb{R}$", "$\mathbb{R}$ **is** a complete ordered field") when we talk about the sets.*

21.1.1 Sets with structure

Most of these collective entities have a ***structure***; that is,

1. The members are not all equal and interchangeable. On the contrary, some play special roles.
2. The pairs of members are not all equal and interchangeable. On the contrary, some pairs of members are different from others.
3. The triples of members are not all equal and interchangeable. On the contrary, some triples of members are different from others.

For example,

1. A flock of birds flying in formation has a special member, the ***leader***. And, even more importantly, each bird has ***neighbors***, that is, a few other birds that are right next to it, to the left or to the right or in front or behind, and the bird communicates with its neighbors. The flock stays in formation because each bird, knowing which way its neighbors are moving, tries to move in the same way. "Being neighbors" is what we have called in these notes a ***binary relation***. If we use "xNy" for "x is a neighbor of y", then the "neighbor" relation N singles out some pairs (x, y) of birds as different from other pairs.
2. A number system such as $\mathbb{N}$, or $\mathbb{Z}$, or $\mathbb{R}$ has
   - special members (1 for $\mathbb{N}$, 0 and 1 for $\mathbb{Z}$ and $\mathbb{R}$),
   - special sets of members (for example, for $\mathbb{Z}$ or for $\mathbb{R}$, the set of all positive members of the set),
   - special pairs of members of the set (for example, for $\mathbb{N}$, $\mathbb{Z}$, or $\mathbb{R}$, the pairs (x, y) such that $x < y$ are different from the other pairs),

- special triples (x, y, z) of members. (For example, the triples (x, y, z) such that $z = x + y$ play a special role: they determine the operation of ***addition***, in the sense that if you know the set S of all the triples (x, y, z) such that $x + y = z$ then you know the operation of addition, because, if I give you numbers x, y, then you can compute x, y by looking in the set S until you find a triple (x, y, z) that is in S, and then the sum $x + y$ is z.)

21.1.2 How sets are different from other collective entities

Usually, you cannot form collective entities by putting together any objects you want, because the objects have to be related in some way. For example,

- You would never form a "crowd" consisting of yourself, the prime minister of Australia, and five people living in Wyoming.
- And you would never take a bunch of wolves living in Wyoming together with some other wolves who live in Sweden and call that a "pack". To form a pack, the wolves have to be together, run together, and hunt together.

Sets are different, in that they are collective entities that can be formed to put together into a single object ***any objects you want***. The things you put together to form a set do not have to be related in any way. For example,

1. You can form a set whose members are all the wolves in Wyoming.
2. You can form a set whose members are all the wolves in Wyoming together with all the wolves in Sweden.
3. You can form a set whose members are three wolves you like who live in Wyoming, together with the musicians of the New York Philharmonic, your uncle Billy, the planets Earth, Mars and Jupiter, the numbers 5, 7 and 23, the numbers π and $3 + \sqrt{5}$, and all the integers that are larger than 377.

The only thing you need in order to be able to form a set S, is a "membership criterion", i.e., a sentence $C(x)$ that specifies the condition that an object x has to satisfy in order to qualify as a member of the set. And any sentence will do[110].

[110] At least for now. Later we will se that we cannot allow absolutely any sentence, because if you do allow that serious trouble ensues, in the form of the "Russell paradox". So we will have to put some limitations. But we are not there yet.

21.1.3 Terms and sentences with variables: a review

In mathematical writing, there are two kinds of meaningful phrases[111], namely, ***terms*** and ***sentences***.

- Terms are phrases that stand for things or people: for example, "Obama", "Alice", "Ronald Reagan", "the table", "the case where I put my sunglasses yesterday", "$2 + 3$", are terms, because they stand for specific things.[112]

- Sentences are phrases that make an assertion that can be true or false: for example, "cows eat grass", "I have no idea where I left the case where I put my sunglasses yesterday", "the planets move around the Sun", "cows like to attack lions and fight them to death", "$2+3 = 5$", "$2+3 = 6$, and "every odd number is prime") are sentences. (Actually, "cows eat grass" is true, "the planets move around the Sun" is true, "cows like to attack lions and fight them to death" is false, "$2+3 = 5$" is true, "$2+3 = 6$" is false, and "every odd number is prime" is true.)

Remark 26. Terms are basically the same as "noun phrases", that is, phrases that can serve as the subject of an "is" sentence. So, for example,

- In the sentence "$2 + 3$ is an odd number", the subject is "$2 + 2$", so "$2 + 2$" is a term.

- In the sentence "the case where I put my sunglasses yesterday is on the table", the subject is "the case where I put my sunglasses yesterday", so "the case where I put my sunglasses yesterday" is a term. □

Terms and sentence can contain variables, that is, letters or expressions that do not stand for a definite object, but represent ***slots*** where the name of a person or object can be inserted. Then, when you actually put specific names of persons or objects in the slots,

[111] A "phrase" in a particular language is, according to the dictionary, "a small group of words standing together as a conceptual unit". (The "small group" could be just a single word. Most phrases are meaningless. For example, the words "Obama" and "Alice" and the longer phrases "Ronald Reagan", "the table", "the case where I put my sunglasses yesterday", "cows eat grass", "the planets move around the Sun", "cows like to attack lions and fight them to death", "$2 + 3$", "$2 + 3 = 5$", "$2 + 3 = 6$", "every odd number is prime", are all phrases.

[112] These things may be concrete,material objects or people, or abstract entities such as numbers. For example, "$2 + 3$" stands for a number, that happens to be the number 5.

- A term has a ***value***, i.e., becomes the name of a specific object.
- A sentence has a ***truth value***, i.e., becomes true or false.

But if you leave some of the the slots unfilled (i.e., if you keep some "free variables") then the terms do not have a definite value and the sentences do not have a truth value. In that case, we say that he term or sentence is meaningless, because it does not stand for a specific object or assertion.

Example 96. The term (i.e., noun phrase) "his mother" contains the possessive adjective "his", which is a variable. If you plug in "Barack Obama" for "his" the term becomes "Barack Obama's mother", which stands for a definite person. (In mathematical language, we would talk about "x's mother". And, again, when we plug in "Barack Obama" for "x" the term becomes "Barack Obama's mother", which stands for a definite person.) □

Example 97. The sentence "he is a friend of mine" contains the pronoun "he". If you do not tell me who "he" is, then I don't know what you are talking about. But if you tell me who "he" is, that is, if you ***assign a value*** to the variable "he" (by saying, for example, that "he" stands for "Bill Clinton") then the sentence becomes "Bill Clinton is a friend of mine", which has a definite truth value. (In mathematical language, we would say "x is a friend of mine", and then, when we plug in "Bill Clinton" for "x", we get when we plug in "Barack Obama" for "x" the term becomes "Barack Obama's mother", which stands for a definite person.) □

Example 98. The term "$x + 3y$" contains the letters "x" and "y". If you do not tell me which numbers the letters x and y stand for, then I cannot make sense of which object (in this case, a number) this term stands for. If, on the other hand, you assign specific values to x and y then I can figure out the value of the term. (For example, if you let $x = 4$, $y = -6$, then I can tell that "$x + 3y$" has the value -14, i.e., that $x + 3y = -14$. □

Example 99. The sentence "$x + 3y > 6$" contains the letters "x" and "y". If you do not tell me which numbers the letters x and y stand for, then I cannot make sense of which assertion the sentence is making, and cannot decide if it is true or false, If, on the other hand, you assign specific values to x and y then I can figure out the truth value of the sentence. (for example, if you let $x = 4$, $y = -6$, then I can tell that "$x + 3y = 6$" has the truth value "false", because $x + 3y = 4 - 3 \times 6 = -14$, and $\sim -14 > 6$. But if $x = 3$ and $y = 2$, then $x + 3y = 9$, and $9 > 6$, so "$x + 3y = 6$" is true.. □

21.1.4 Forming sets

As long as you can write a sentence $C(x)$ about a variable object x, you can form the set

$$\{x : C(x)\}$$

that is, the set of all x for which $C(x)$ is true. And you could give this set a name. For example, suppose you want to form the set $\{x : C(x)\}$ and give it the name S. You would do that by writing

$$\text{Let} \qquad S = \{x : C(x)\}.$$

Let us formulate this rule for forming sets as an axiom:

The naïve axiom of set formation

Given any sentence $C(x)$ having x as an open variable, we can form the set whose members are all the objects x for which $C(x)$ is true.

A name for such a set is

$$\{x : C(x)\}\,.$$

And we read this as

The set of all x such that $C(x)$.

Remark 27. Why did I call the set formation axiom "naïve"? The reason is this: in a few days, we will discover that the set formation axiom, as we have formulated it, causes serious problems that can only be solved by changing the statement of the axiom. Instead of a "naïve" axiom that allows us to take any sentence $C(x)$ whatsovever and form the ser $\{x : C(x)\}$, we will have to adopt a "sophisticated" axiom in which nto all sentences are permitted. □

21.1.5 The membership criterion

Suppose we use the sentence "x is a cow", to form a set S, so

$$S = \{\, x : x \text{ is a cow} \,\}$$

that is, S is "the set of all x such that x is a cow", or, in much better English, S ***is the set of all cows.***

Then we can decide whether or not an object a belongs to the set S (that is, whether or not $a \in S$) by applying the following simple test

1. Find out if a is a cow or not.
2. If a is a cow, then a belongs to S.
3. If a is not a cow, then a does not belong to S.

In other words, the sentence "x is a cow" is the ***membership criterion***, or ***membership condition***, for S. A particular object a belongs to the set $\{\, x : x \text{ is a cow} \,\}$ if a is a cow, and doesn't belong to the set of a is not a cow.

For a general sentence $C(x)$:

Suppose $C(x)$ is a sentence having x as an open variable, and you define a set S by writing

$$\text{Let} \qquad S = \{x : C(x)\}\,.$$

Then

- The sentence $C(x)$ is called the membership criterion, or membership condition, for the set S.
- An object a ***belongs*** to S if $C(a)$ is true, and ***doesn't belong*** to S if $C(a)$ is not true.

21.1.6 Forming sets of members of a given set

Suppose we want to form the set of all natural numbers n that are even, i.e., such that $2|n$, and we want to call this set A.

Then we can say:

$$\text{Let} \qquad A = \{n : n \in \mathbb{N} \wedge 2|n\}\,,$$

and we can also say

$$\text{Let} \qquad A = \{n \in \mathbb{N} : 2|n\}\,.$$

The first ssentence is read as "Let A be the set of all things that are natural numbers and are even", whereas the second sentence is read as "Let A be the set of all natural numbers that are even".

And, clearly, both define the same set.

Suppose U is a set, $C(x)$ is a sentence having x as an open variable, and you define a set S by writing

$$\text{Let} \qquad S = \{\, x : x \in U \wedge C(x) \,\}\,.$$

Then the membership criterion is the sentence "$x \in U \wedge C(x)$".
And you can also write

$$\text{Let} \qquad S = \{\, x \in U : C(x) \,\}\,.$$

Example 100. Suppose the membership criterion $C(x)$ is the sentence "x is a natural number that can be written as the sum of the squares of two natural numbers". Let

$$S = \{x : C(x)\}\,.$$

Clearly, $C(x)$ is the sentence

$$x \in \mathbb{N} \wedge (\exists m \in \mathbb{N})(\exists n \in \mathbb{N})x = m^2 + n^2\,,$$

so we could have written the definition of S as follows:

$$S = \{x : x \in \mathbb{N} \wedge (\exists m \in \mathbb{N})(\exists n \in \mathbb{N})x = m^2 + n^2\}\,,$$

or as

$$S = \{x \in \mathbb{N} : (\exists m \in \mathbb{N})(\exists n \in \mathbb{N})x = m^2 + n^2\}\,, \tag{21.467}$$

(We read this as "S is the set of all natural numbers x such that there exist natural numbers m, n for which $m^2 + n^2 = x$". And an even better reading is "S is the set of all natural numbers that are the sum of two squares of natural numbers".)

Let us consider several possible values of x, and in each case let us figure out whether this x belongs to the set S.

1. Suppose x is the Math 300 textbook. Then x is a book, not a natural number. So $x \notin S$, that is, x is not a member of S.

2. Suppose $x = 5$. Then x is a natural number. And x is the sum of the squares of two natural numbers, because $x = 2^2 + 1^2$. Therefore x satisfies the criterion for membership in S. So x is a member of S, that is, $x \in S$.

3. Suppose $x = -5$. Then x is not a natural number. So $C(x)$ is not true. That is, x does not satisfy the criterion for membership in S. So x is not a member of S.

4. Suppose $x = 7$. Then x is a natural number. Can x be written as the sum of the squares of two natural numbers? The answer is "no". How do we know that? Well, for example, we know that a number that is of the form $k + 3$, $k \in \mathbb{Z}$, is not the sum of two squares. And 7 is of the form $k + 3$, because $7 = 4 + 3$. So $x \notin S$. □

21.1.7 How to read the symbol "$\in$"

How to read the "$\in$" symbol

If S is a set and a is an object, we write

$$a \in S$$

to indicate that a is a member of S.

And we write

$$a \notin S$$

to indicate that a is not a member of S.
The expression "$a \in S$" is read in any of the following ways:

- a belongs to S,
- a is a member of S,
- a is in S.

The expression "$a \notin S$" is read in any of the following ways:

- a does not belong to S,
- a is not a member of S,
- a is not in S.

Remark 28. Sometimes, "$a \in S$" is read as "a belonging to S", or "a in S", rather than "a belongs to S", or "a is in S." For example, if we write

Pick an $a \in S$,

then it would be very bad to say "pick an a belongs to S". But "pick an a belonging to S", "pick an a in S", is fine. □

Never read "$\in$" as "is contained in", or "is included in". The words "contained" and "included" have different meanings, that will be discussed later.

21.2 When are two sets equal?

As we have explained, sets have ***members***. And, even more imprtantly, ***knowledge of the members of the set determines the set. Two sets that have the same members are the same set.***

Let us make this precise:

The axiom of set equality

Two sets are equal if and only if they have the same members.
In semiformal language:
If A, B are sets, then $A = B$ if and only if

$$(\forall x)(x \in A \iff x \in B)\,.$$

And, in formal language,

$$(\forall A)(\forall B)\Big(A = B \iff (\forall x)(x \in A \iff x \in B)\Big)\,.$$

Example 101. Let

$$\begin{aligned} A &= \{x \in \mathbb{R} : x \geq 0\}\,, \\ B &= \{x \in \mathbb{R} : (\exists y \in \mathbb{R}) y^2 = x\}\,. \end{aligned}$$

Let us prove that $A = B$.
To prove that $A = B$, we have to prove that $(\forall x)(x \in A \iff x \in B)$.

So, let x be arbitrary. We have to prove that $x \in A \iff x \in B$.

To prove this, we have to prove that $x \in A \implies x \in B$ and that $x \in B \implies x \in A$.

Let us first prove that $x \in A \implies x \in B$.

Assume that $x \in A$.

Then $x \in \mathbb{R}$ and $x \geq 0$. (Reason: "$x \in \mathbb{R} \wedge x \geq 0$" is the membership criterion for A.)

But every nonnegative real number has a square root.

So x has a square root. That is, $(\exists y \in \mathbb{R}) y^2 = x$.

So x satisfies the membership criterion for B.

Hence $x \in B$.

Therefore $x \in A \Longrightarrow x \in B$.

We now prove that $x \in B \Longrightarrow x \in A$.

Assume that $x \in B$.

Then $x \in \mathbb{R}$ and $(\exists y \in \mathbb{R}) y^2 = x$. (Reason: "$x \in \mathbb{R} \wedge (\exists y \in \mathbb{R}) y^2 = x$" is the membership criterion for B.)

Pick $y \in \mathbb{R}$ such that $y^2 = x$.

Then $y^2 \geq 0$. (Reason: $(\forall u \in \mathbb{R}) u^2 \geq 0$.)

So $x \geq 0$.

So x satisfies the membership criterion for A.

Hence $x \in A$.

Therefore $x \in B \Longrightarrow x \in A$.

So $x \in A \Longleftrightarrow x \in B$. Since x is arbitrary, we can conclude that $(\forall x)(x \in A \Longleftrightarrow x \in B)$. Hence $A = B$. **Q.E.D.**

Example 102. Let

$$\begin{aligned} A &= \{x \in \mathbb{R} : x > 0\}, \\ B &= \{x \in \mathbb{R} : (\exists y \in \mathbb{R}) y^2 = x\}. \end{aligned}$$

Let us prove that $A \neq B$.

To prove that $A \neq B$, we have to prove that it is not true that $(\forall x)(x \in A \Longleftrightarrow x \in B)$.

Suppose[113] $(\forall x)(x \in A \Longleftrightarrow x \in B)$.

Then we can specialize to $x = 0$, and conclude that $0 \in A \Longleftrightarrow 0 \in B$.

[113] A proof by contradiction , of course.

But "$0 \in B$" means that "$(\exists y \in \mathbb{R}) y^2 = 0$, which is true, because $7^2 = 0$.

On the other hand, "$0 \in A$" means that "$0 > 0$", which is false.

Hence it is not true that $0 \in A \iff 0 \in B$.

So $(0 \in A \iff 0 \in B) \wedge \Big(\sim (0 \in A \iff 0 \in B)\Big)$, which is a contradiction .

Hence $A \neq B$. **Q.E.D.**

Example 103. Let $A = \{n \in \mathbb{Z} : 6|n\}$, and let $B = \{n \in \mathbb{Z} : 2|n \wedge 3|n\}$.
Let us prove that $A = B$.
To prove that $A = B$, we have to prove that $(\forall x)(x \in A \iff x \in B)$.

So, let x be arbitrary. We have to prove that $x \in A \iff x \in B$.

To prove this, we have to prove that $x \in A \implies x \in B$ and that $x \in B \implies x \in A$.

Let us first prove that $x \in A \implies x \in B$.

Assume that $x \in A$.
Then $x \in \mathbb{Z}$ and $6|x$.
Since $6|x$, we may pick $k \in \mathbb{Z}$ such that $x = 6k$.
Then $x = 2 \times (3k)$, and $3k \in \mathbb{Z}$, so $2|x$.
Also, $x = 3 \times (2k)$, and $2k \in \mathbb{Z}$, so $3|x$.
Hence $2|x \wedge 3|x$.
So $x \in B$.

Therefore $x \in A \implies x \in B$.

We now prove that $x \in B \implies x \in A$.

Assume that $x \in B$.
Then $x \in \mathbb{Z}$, $2|x$, and $3|x$.
Since $2|x$, we may pick $j \in \mathbb{Z}$ such that $x = 2j$.
Since $3|x$, we may pick $k \in \mathbb{Z}$ such that $x = 3k$.
Then $x = 1.x = (3-2)x = 3x-2x = 3\times(2j)-2\times(3k) = 6(j-k)$.
So $6|x$.

Hence $x \in A$.

Therefore $x \in B \Longrightarrow x \in A$.

So $x \in A \Longleftrightarrow x \in B$. Since x is arbitrary, we can conclude that $(\forall x)(x \in A \Longleftrightarrow x \in B)$. Hence $A = B$. **Q.E.D.**

Problem 100. Let

$$\begin{aligned} A &= \{x \in \mathbb{R} : x^3 > x\}, \\ B &= \{x \in \mathbb{R} : -1 < x < 0 \vee x > 1\} \\ C &= \{x \in \mathbb{R} : -1 < x\}. \end{aligned}$$

Prove or disprove each of the following:

- $A = B$,
- $A = C$.

21.2.1 Subsets

Definition 41. *Let A, B be sets. We say that A is a <u>subset</u> of B, and write*

$$A \subseteq B,$$

if every member of A is a member of B.

In semiformal language, A is a subset of B if and only if

$$(\forall x)(x \in A \Longrightarrow x \in B).$$

In completely formal language:

$$\boxed{(\forall A)(\forall B)\Big(A \subseteq B \Longleftrightarrow (\forall x)(x \in A \Longrightarrow x \in B)\Big).}$$

□

Example 104. The following are true:

- $\mathbb{N} \subseteq \mathbb{Z}$,

- $\mathbb{Z} \subseteq \mathbb{Q}$,
- $\mathbb{Q} \subseteq \mathbb{R}$,
- $\{x \in \mathbb{R} : 0 < x < 1\} \subseteq \{x \in \mathbb{R} : 0 \leq x \leq 1\}$. □

Example 105.

The following are true:

- $\{x \in \mathbb{R} : -1 < x < 0\} \subseteq \{x \in \mathbb{R} : x^3 > x\}$.
- $\{n \in \mathbb{N} : n \text{ is prime} \wedge n \neq 2\} \subseteq \{n \in \mathbb{N} : 2|n-1\}$.
- $\{n \in \mathbb{Z} : 4|n\} \subseteq \{n \in \mathbb{Z} : 2|n\}$,
- $\{x \in \mathbb{R} : 0 < x < 1\} \subseteq \{x \in \mathbb{R} : 0 \leq x \leq 1\}$, □

WARNING!

"is a subset of" is a ***binary relation***. It does not make sense to say things like "A is a subset". What does make sense is to say "A is a subset of B".
If, in an exam, I ask you to define "subset", and you say "a set A is a subset if ....", then that is completely wrong and you get zero credit[a]
The definition of "subset" must start with the words: "Let A, B be sets. We say that A is a subset of B if ....

[a] And if your definition starts with horrendous words "subset is when ..." then you lose $10,000,000$ points, on a sclae from 0 to 10.

ALWAYS UNDERLINE THE DEFINIENDUM

In a definition, the term being defined is called the definiendum. The definiendum must always be underlined, or highlighted in some way, in order to indicate that we are writing a definition of that term, not just making a true statement.
For example:

- If I write "elephants are four-legged animals", then I am making a true statement about elephants.
- If, on the other hand, I write "elephants are four-legged animals", then I am saying that I am defining the word "elephant" to mean "four-legged animal", and this is of course wrong, because "elephant" does not mean "four-legged animal": there are lots of four-legged animals that are not elephants.
- If I write "an even integer is an integer that is divisible by 2", then I am making a true statement. but I am not saying that this is what "even integer" means.
- If I want to explain what "even integer" means, i.e., give a ***definition*** of "even integer", then I have to say "an even integer is an integer that is divisible by 2". By underlining "even integer" I am conveying the message that this is my definition of "even inteeger".
- If in an exam you are asked to give a definition and you do not underline the definiendum, you will lose points.

Question 10. *In the first sentence of the previous box, why is the word "definiendum" underlined?* □

Problem 101. ***Prove*** the four statements of Example 105.
The structure of your proofs should be as follows:

We want to prove that $A \subseteq B$.

For that purpose, we prove that $(\forall x)(x \in A \Longrightarrow x \in B)$.

Let x be arbitrary. We want to prove "$x \in A \Longrightarrow x \in B$".

Assume $x \in A$.
$\vdots$
$x \in B$.
So $x \in A \Longrightarrow x \in B$.

Therefore $(\forall x)(x \in A \Longrightarrow x \in B)$.

So $A \subseteq B$. **Q.E.D.**

Problem 102. ***Prove*** that the binary relation "$\subseteq$" is reflexive, antisymmetric, and transitive. (In the definition of these properties given in the notes, a set S is mentioned. Here you may think of S as "the set of all sets", which means that you can forget about S. Then, for example, the property that "$\subseteq$" is antisymmetric means "$(\forall A)(\forall B)\Big((A \subseteq B \wedge B \subseteq A) \Longrightarrow A = B\Big)$".)

21.2.2 The empty set

An important example of a set is the ***empty set***, that is, the set that has no members at all.

The symbol for the empty set is

$$\emptyset\,.$$

One possible way to define this set is by the following formula:

$$\boxed{\emptyset = \{x : x \neq x\}\,.}$$

This means that the members of $\emptyset$ are the things x that satisfy $x \neq x$. But our Equality Axiom says that $(\forall x)x = x$. So "$x = x$" is true for every x. This means that no x can be a member of $\emptyset$. So, indeed. $\emptyset$ has no members.

Let us make this precise:

Theorem 85. *The empty set has no members. That is.*

$$(\forall x) x \notin \emptyset \,.$$

Proof.

Let x be arbitrary. We want to prove that $x \notin \emptyset$.

Assume[114] that $x \in \emptyset$.

Then x satisfies the membership criterion for $\emptyset$, i.e.,

$$x \neq x \,.$$

But $(\forall x) x = x$, by the Equality Axiom.

So $x = x$, by the rule for using universal sentences.

Therefore $x = x \wedge x \neq x$, which is a contradiction.

So $x \notin \emptyset$.

Therefore $(\forall x) x \notin \emptyset$. **Q.E.D.**

21.2.3 The empty set is a subset of every set

If you have a set A and a subset B of A, and you remove some members from B, producing a subset C of B, then it is clear that C is still a subset of A. This ought to be true even in the extreme case when you remove *all* the members of B, so that C is the empty set. In other words, the empty set should be a subset of A, for every set A.

Let us prove a precise theorem:

Theorem 86. *The empty set is a subset of every set. That is,*

$$(\forall A) \emptyset \subseteq A \,.$$

Proof.

Let A be an arbitrary set. We want to prove that $\emptyset \subseteq A$.

Assume[115] that $\emptyset$ is not a subset of A.

That is, assume that it is not true that every member of $\emptyset$ is in A.

[114] A proof by contradiction !.
[115] A proof by contradiction !

That means that some members of $\emptyset$ are not in A.

In other words, there exists an object x such that $x \in \emptyset$ and $x \notin A$.

Pick one such object and call it a.

Then $a \in \emptyset$ and $a \notin A$.

So in particular $a \in \emptyset$.

But we know from Theorem 85 that $(\forall x) x \notin \emptyset$.

So $a \notin \emptyset$.

Hence $a \in \emptyset \wedge a \notin \emptyset$.

So we have proved a contradiction.

Therefore $\emptyset \subseteq A$.

So $(\forall A)\emptyset \subseteq A$. **Q.E.D.**

21.2.4 Sets with one, two, three or four members

If a is any thing, we can form a set that has a as a member, and no other members. This name of this set is

$$\boxed{\{a\}},$$

which we read as "singleton of a."

The precise definition of $\{a\}$ is as follows.

Definition 42. *Let a be any object. Then the singleton of a is the set $\{a\}$ given by*

$$\{a\} = \{x : x = a\}.$$

In other words: to be a member of the set $\{a\}$ you have to be a. If you are a then you are a member, and if you are not a then you are not a member.

We can do a similar thing with two objects, say a and b. We can form the set $\{a, b\}$ whose members are a, b, and nothing else. The set $\{a, b\}$ is the unordered pair of a and b.

Definition 43. *Let a, b be any two objects. Then the unordered pair of a and b is the set $\{a, b\}$ given by*

$$\{a, b\} = \{x : x = a \vee x = b\}.$$

Remark 29. ***Warning:*** The set $\{a, b\}$ is ***not*** necessarily a set with two members. That depends on who a and b are. For example; if a happens to be equal to b, then $\{a, b\}$ has only one member. □

Naturally, we can do the same thing with three, four, or any number of objects. For example:

Definition 44. *Let a, b, c be any three objects. Then the unordered triple of a, b and c is the set $\{a, b, c\}$ given by*

$$\{a, b, c\} = \{x : x = a \vee x = b \vee x = c\}.$$

Definition 45. *Let a, b, c, d be any four objects.*

Then the unordered quadruple of a, b, c and d is the set $\{a, b, c, d\}$ given by

$$\{a, b, c, d\} = \{x : x = a \vee x = b \vee x = c \vee x = d\}.$$

And, in principle, you could go on like this and define sets with five members, sets with 6 members, and so on.

But as soon as the number of members gets large, this way of constructing sets becomes very complicated, so it is better to do it differently.

Example 106. Suppose you want to define a set whose members are the first five presidents of the U.S., and call this set A. That's easy to do. We say:

Let $A = \{$George Washington,John Adams,Thomas Jefferson,James Madison,James Monroe$\}$.

Now suppose you want to define a set whose members are the first 30 U.S. presidents, and call this set B. That is going to be much more complicated right? And what if you do not know the names of all those presidents?

Hhere is how you can do it. You can say:

Let

$$B = \left\{ x : (\exists j \in \mathbb{N})(j \leq 30 \wedge x = p_j) \right\},$$

where, for each $j \in \mathbb{N}$, p_j is the j-th president of the U.S.

This works perfectly! Indeed, let us see what has to be true of an object x for x to qualify as a member of A. If you are given an object x, and you

have to decide whether $x \in B$ or not, you have to find out if there exists a natural number j such that $j \leq 30$ and x is the j-th U.S. president. And that's exactly what we want! □

Problem 103. How many members does the set B of Example 106 have?

If you think that the answer is 30, think again! Go to a history book (or to a history Web site) and read about Grover Cleveland, who was both the 22nd and the 24th president of the United States. □

Problem 104. Let $A = \{1, 2, 3, 4\}$. Write a list of all the subsets of A. (HINT: There are 16 of them.) □

Problem 105. Write a definition, in the style of Example 106, of the set X whose members are the first 325 prime numbers p such that $p-3$ is divisible by 4. □

21.3 Operations on sets

There are several operations that enable us to construct new sets from given sets.

21.3.1 The power set of a set

Definition 46. *Let A be a set. The power set of A is the set $\mathcal{P}(A)$ given by*

$$\mathcal{P}(A) = \{X : X \subseteq A\}.$$

In other words, $\mathcal{P}(A)$ ***(read as "the power set of A") is the set whose members are all the subsets of A.***

The ***membership criterion*** for the power set $\mathcal{P}(A)$ is the sentence "$X \subseteq A$". That is, for an object X to quality as a member of $\mathcal{P}(A)$, it has to be shown that X is a subset of A.

Example 107. If $A = \{1, 2, 3\}$ then

$$\mathcal{P}(A) = \Big\{ \emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\} \Big\}. \tag{21.468}$$

Notice that A is a finite set with 3 members, and $\mathcal{P}(A)$ has turned out to be a finite set with 8 members. ***This is not a coincidence. We will***

prove later that: if A ***is a finite set and*** A ***has*** n ***members, then the power set*** $\mathcal{P}(A)$ ***is a finite set with*** 2^n ***members.*** □

Problem 106. Let $A = \{1, 2, 3, 4\}$. Write a formula similar to (21.468) listing all the members of $\mathcal{P}(A)$.

Problem 107. Let $A = \{\emptyset, \{\emptyset\}\}$. Write a formula similar to (21.468) listing all the members of $\mathcal{P}\Big(\mathcal{P}(A)\Big)$.

21.3.2 The union of two sets

Definition 47. *Let A, B be sets. The union of A and B is the set $A \cup B$ given by*

$$A \cup B = \{x : x \in A \vee x \in B\}.$$

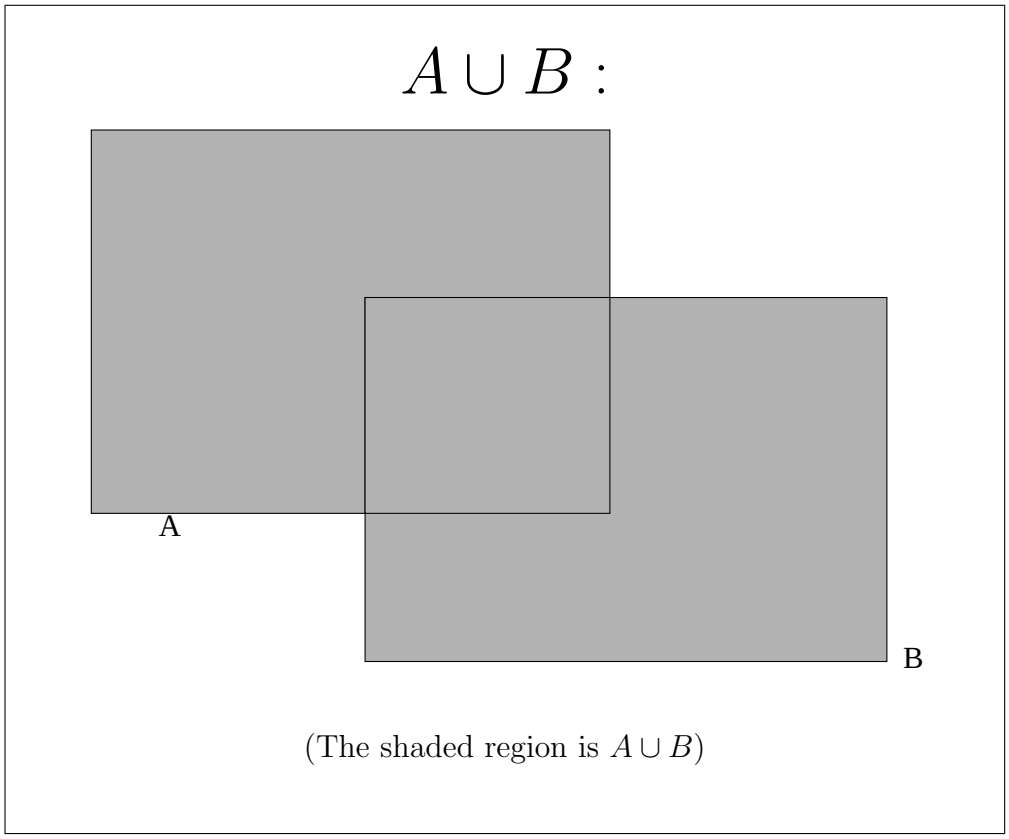
In other words, $A \cup B$ ***(read as "*** A ***union*** B***") is the set whose members are all the members of*** A ***as well as all the members of*** B***.***

The ***membership criterion*** for $A \cup B$ is "$x \in A \vee x \in B$." That is, for an object x to quality as a member of $A \cup B$, it has to be shown that x is in A or that x is in B.

Example 108.

- If $A = \{1, 2, 3\}$ and $B = \{2, 3, 4\}$ then $A \cup B = \{1, 2, 3, 4\}$.
- If $A = \{a, b, c\}$ and $B = \{d, e, f, g, h, i, j\}$ then $A \cup B = \{a, b, c, d, e, f, g, h, i, j\}$. *Notice that*
  1. *A is a finite set with 3 members,*
  2. *B is a finite set with 7 members,*
  3. *A and B have no memebrs in common (that is, using the terminology of the next section, $A \cap B = \emptyset$),*
  4. *and $A \cup B$ has turned out to be a finite set with 10 members.* ***This is not a coincidence. We will prove later that: if*** A***,*** B ***are finite sets,*** A ***has*** m ***members,*** B ***has*** n ***members, and*** $A \cap B = \emptyset$***, then the union*** $A \cup B$ ***is a finite set with*** $m + n$ ***members.***

5. *If* $A = \{n \in \mathbb{Z} : n > 0\}$ *and* $B = \{n \in \mathbb{Z} : n < 0\}$ *then* $A \cup B = \{n \in \mathbb{Z} : n \neq 0\}$.
6. $\mathbb{N} \cup \{0\}$ *is the set of all nonnegative integers, i.e., the set* $\{n \in \mathbb{Z} : n \geq 0\}$.
7. *If* $A = \{x \in \mathbb{R} : 0 < x < 1\}$ *and* $B = \{x \in \mathbb{R} : 1 \leq x < 2\}$ *then* $A \cup B = \{x \in \mathbb{R} : 0 < x < 2\}$.
8. *If* $A = \{x \in \mathbb{R} : 0 < x < 1\}$ *and* $B = \{x \in \mathbb{R} : 1 < x < 2\}$ *then* $A \cup B = \{x \in \mathbb{R} : 0 < x < 2 \wedge x \neq 1\}$. □



21.3.3 The intersection of two sets

Definition 48. *Let A, B be sets. The intersection of A and B is the set $A \cap B$ given by*

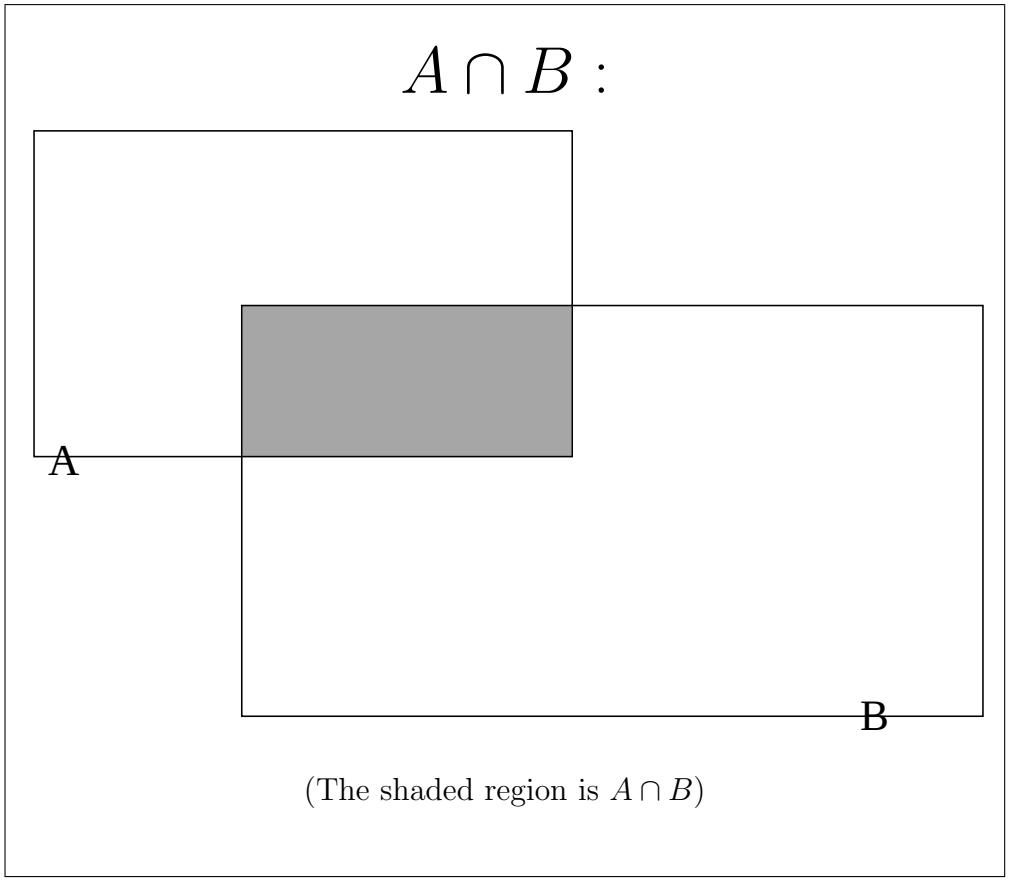
$$A \cap B = \{x : x \in A \wedge x \in B\}.$$

In other words, $A \cap B$ ***(read as "A intersection B") is the set whose members are all the things that belong both to A and to B.***

The ***membership criterion*** for $A \cap B$ is "$x \in A \wedge x \in B$." That is, for an object x to quality as a member of $A \cap B$, it has to be shown that x is in A and that x is in B.

Example 109.

- If $A = \{1, 2, 3\}$ and $B = \{2, 3, 4\}$ then $A \cap B = \{2, 3\}$.
- If $A = \{n \in \mathbb{Z} : n > 0\}$ and $B = \{n \in \mathbb{Z} : n < 0\}$ then $A \cap B = \emptyset$.
- If $A = \{x \in \mathbb{R} : 0 < x < 2\}$ and $B = \{x \in \mathbb{R} : 1 < x < 3\}$ then $A \cap B = \{x \in \mathbb{R} : 1 < x < 2\}$.



21.3.4 The difference of two sets

Definition 49. *Let A, B be sets. The difference of A and B is the set $A - B$ given by*

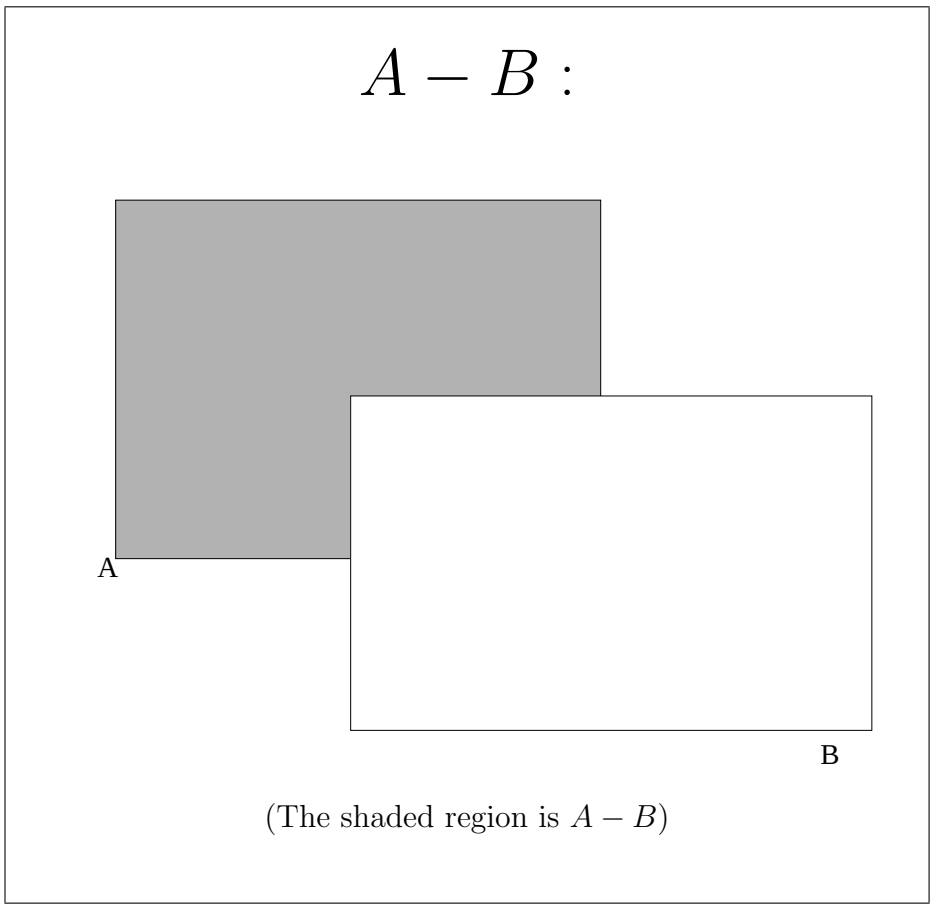
$$A - B = \{x : x \in A \wedge x \notin B\}.$$

In other words, $A - B$ ***(read as "A minus B") is the set whose members are all the things that belong to A but do not belong to B.***

The ***membership criterion*** for $A - B$ is "$x \in A \wedge x \notin B$." That is, for an object x to quality as a member of $A - B$, it has to be shown that x is in A and that x is not in B.

Example 110.

- If $A = \{1, 2, 3\}$ and $B = \{2, 3, 4\}$ then $A - B = \{1\}$.
- If $A = \mathbb{Z}$ and $B = \mathbb{N}$ then $A - B = \{n \in \mathbb{Z} : n \geq 0\}$.
- If $A = \{x \in \mathbb{R} : 0 < x < 2\}$ and $B = \{x \in \mathbb{R} : 1 < x < 3\}$ then $A - B = \{x \in \mathbb{R} : 0 < x \leq 1\}$.



21.3.5 Complements

As you may have noticed, the operations of union and intersection are cloely related to the logical connectives $\vee$ and $\wedge$:

$A \cup B$ is the set of those x such that $x \in A \vee x \in B$
$A \cap B$ is the set of those x such that $x \in A \wedge x \in B$

Given this, which is rhe set operation that corresponds to the negation symbol $\sim$? Since $\sim$ is a unary connective (i.e., it can be applied to one sentence S to produce the sentence $\sim S$. the corresponding operation, let us call it $\#$, should be a unary operation defined as follows:

$\#A$ is the set of those x such that $\sim x \in A$.

In other words, $\#A$ should be the set of all the things that are not members of A. This set $\#A$ could be called the "complement" of A, and would be defined by $\#A = \{x : x \notin A\}$.

Now, the set $\#A$ would be truly huge. For example, if $A = \{1, 2, 3, 4\}$, then $\#A$ would consist of all the things other than the numbers $1, 2, 3, 4$. So the members of $\#A$ would be the natural numbers other than $1, 2, 34$ (that is, $5, 6, 7$ and so on), as well as the integers that are not hantural numbers, all the real numbers other than $1, 2, 3, 4$, plus all the other things that are not the numbers $1, 2, 3, 4$, that is, all the cows, sheep, giraffes, people, rocks, tables, planets, stars, cells, viruses, molecules, atoms, electorns, protons, quarks, black holes, books, teeth, jackets, socks, cars, planes, forks, knives, and on and on and on.

Usually, when we are doing mathematics, we are studying a specific "universe" of mathematical objects. For example, when we do number theory we study the natural numbers or the integers, when we do Calculus we work with the real nunbers, and when we do Multivariable Calculus we work with $\mathbb{R}^2$, the set of pairs of real numbers)(i.e., the "xy plane") or $\mathbb{R}^3$ (the set of triples (x, y, z) of real numbers, i.e., "3-dimensional space"). If, for example, our "world" is $\mathbb{R}$, then when we have a set A of real numbers, i.e., a subset A of $\mathbb{R}$, we would be interested in the set of real numbers that are not in A. And this set is the difference $\mathbb{R} - A$. Se we give the following definition:

Definition 50. *Suppose U is a set that we regard as the "universe", in the sense that we are only interested in sets that are subsets of U. Then the complement of a set A such that $A \subseteq U$ is the set A^c given by*

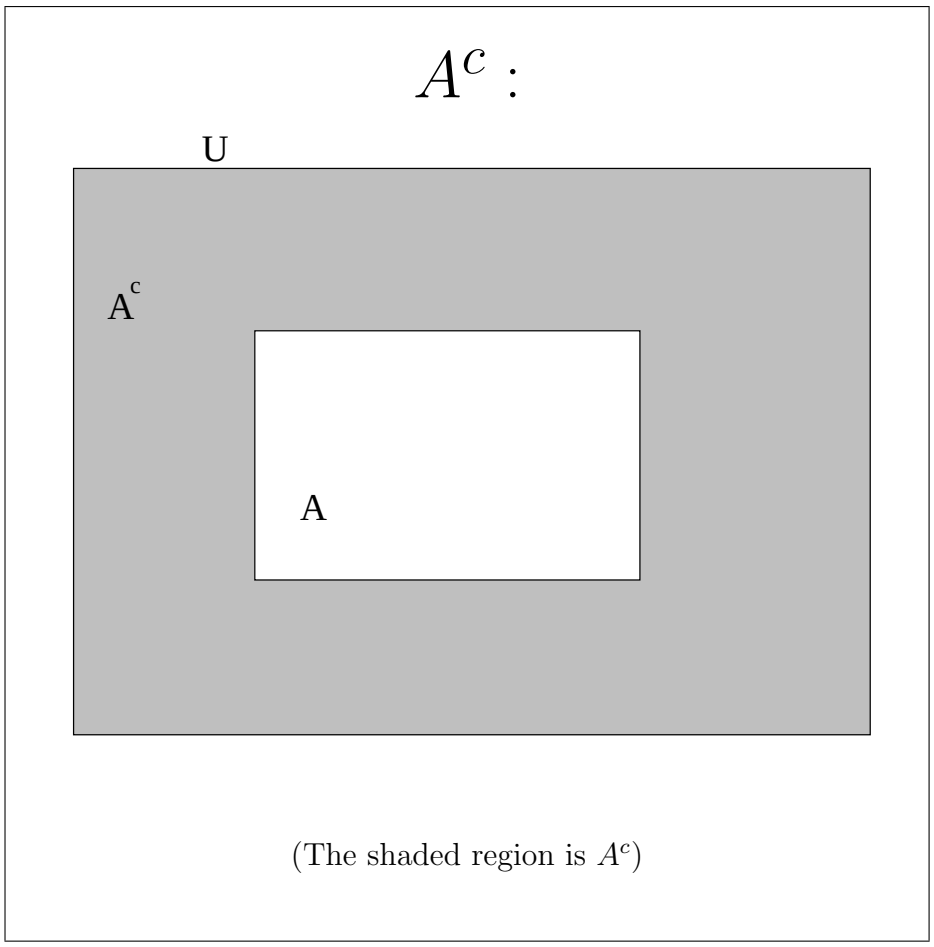
$$A^c = U - A\,, \tag{21.469}$$

that is,

$$A^c = \{x : x \in U \wedge x \notin A\}\,. \tag{21.470}$$

Remark 30. Strictly speaking, it is inappropriate to define a set as we did in Definition 50 and call it "A^c". This set depends very much on who U is, so the right thing to do would be to call it the ***complement of*** A ***relative to*** A, and give it a name such as $A^{c,U}$, which shows that the set depends on U.

But, as long as we are working with a fixed "universe", and it is clear who U is, it is O.K. to use a notation such as A^c. □



21.3.6 The symmetric difference of two sets

Definition 51. *Let A, B be sets. The symmetric difference of A and B is the set $A\Delta B$ given by*

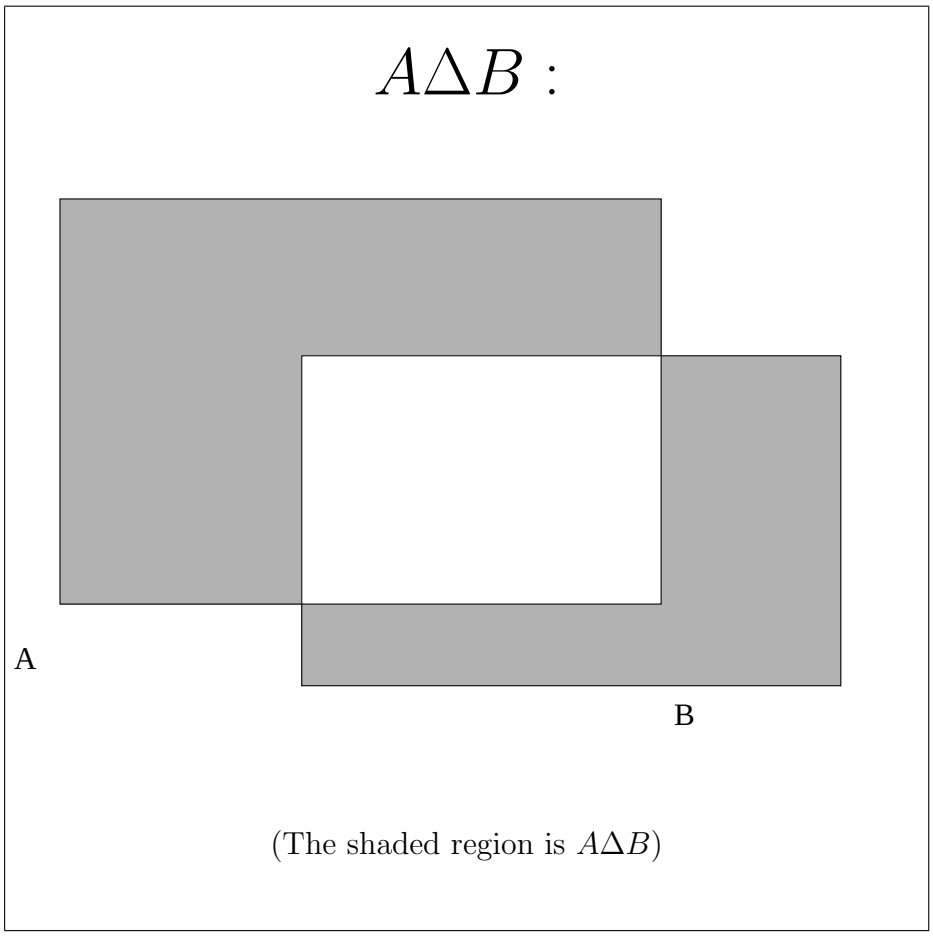
$$A\Delta B = \{x : (x \in A \wedge x \notin B) \vee (x \notin A \wedge x \in B)\}.$$

In other words, $A\Delta B$ ***(read as "the symmatric difference of** A **and***

B") is the set whose members are all the things that belong to A but do not belong to B, or belong to B but do not belomng to A.

That is, $A\Delta B$ ***is the set of all things that belong to one of the sets A, B but do not belong to both.***

The ***membership criterion*** for the symmetric difference $A\Delta B$ is the sentence "$(x \in A \wedge x \notin B) \vee (x \notin A \wedge x \in B)$". That is, for an object x to quality as a member of $A - B$, it has to be shown that x is in A and that x is not in B, or that x is in B but not in A.



Example 111.

- If $A = \{1, 2, 3\}$ and $B = \{2, 3, 4\}$ then $A\Delta B = \{1, 4\}$.
- If $A = \{x \in \mathbb{R} : |x| > 4\}$ and $B = \{x \in \mathbb{R} : |x| < 10$ then $A\Delta B = \{x \in \mathbb{R} : |x| \geq 10 \vee |x| \leq 4\}$.

21.4 Orderd pairs and Cartesian products

21.4.1 Ordered pairs

If a, b are any two objects, we would like to have a set, called "the ordered pair of a and b", such that wknowing this set would tell us who a is and who b is, so we would be able to say things such as "the first coordinate of (a, b) is a" and "the second coordinate of (a, b) is b".

For example, suppose we are doing plane geometry, using the standard procedure of drawing and "x axis" and a "y axis", and then representing each point P of the plane by a pair (a, b) of numbers, called the "coordinate pair" of P. Each point P then has, attached to it, a coordinate pair (a, b) of real numbers: the number a is the x ***coordinate*** (or "abscissa") and the number b is the y ***coordinate*** (or "ordinate") of P.

We would like the pair (a, b) to be a set, constructed somehow from a and b. And then the natural question is: which set is the pair (a, b)?

The most naïve idea is to let the pair (a, b) be the unordered pair $\{a, b\}$, that is, the set whose members are a and b.

But this will not do. If we take (a, b) to be $\{a, b\}$, then it cannot happen, for example, that the x-coordinate of $(1, 2)$ is 1, and the x-coordinate of $(2, 1)$ is 2, because, if $(1, 2) = \{1, 2\}$ and $(2, 1) = \{2, 1\}$, then $(1, 2) = (2, 1)$, so, if

(*) the x-coordinate of $(2, 1)$ is 2,

then it would also be true that

(**) the x-coordinate of $(1, 2)$ is 2,

(because $(1, 2) = (2, 1)$), but on the other hand

(***) the x-coordinate of $(1, 2)$ is 1,

so we would get $1 = 2$, which is definitely not true.

The only solution is to define the ordered pair to be something other than the unordered pair $\{a, b\}$. And then the question is, ***what set shall*** (a, b) ***be?***

There are many ways to answer this question, and it really makes no difference which one we use. So we shall choose one, but you must be warned that the specific way we make this choce is not important. What is important is that the following fact is true:

Theorem 87. *Let a, b, c, d be any objects. Then, if the pairs (a, b) and (c, d) are equal, that is, if $(a, b) = (c, d)$, it follows that $c = a$ and $d = b$.*

This is exactly the property that we need. For example, the pairs $(2, 1)$ and $(1, 2)$ are ***not*** equal. (Proof: Suppose $(2, 1) = (1, 2)$. Then Theorem 87

(with $a = 2$, $b = 1$, $c = 1$, and $d = 2$, would imply that $2 = 1$. But $2 \neq 1$. So $2 = 1 \wedge 2 \neq 1$, which is a contradiction. So $(2, 1) \neq (1, 2)$.)

Now we show how to define (a, b) in such a way that Theorem 87 is true.

Definition 52. *Let a, b be any two objects. Then the ordered pair of a and b is the set (a, b) given by*

$$(a, b) = \{\, \{a\}\,,\, \{a, b\}\,\}\,. \tag{21.471}$$

Proof of Theorem 87. Suppose that $(a, b) = (c, d)$.
Let $p = (a, b)$, so p is also equal to (c, d) because we are assuming that $(a, b) = (c, d)$.
Since $p = \{\, \{a\}, \{a, b\}\,\}$, the set p has either two members (if $b \neq a$) or one member (if $a = b$, in which case $\{a, b\} = \{a\}$, so $\{\, \{a\}, \{a, b\}\,\} = \{\, \{a\}, \{a\}\,\} = \{\, \{a\}\,\}$).
But in either case, a is the only object that belongs to all the members of p.
And, since p is also equal to (c, d), it follows that c is the only object that belongs to all the members of p.
So $\boxed{\boxed{c = a}}$.
Next, let us prove that $d = b$.
We consider separately the two possible cases: $b = a$ and $b \neq a$.

> Assume that $b = a$.
>
> Then p has only one member, because, as explained before, $\{a, b\} = \{a\}$, so $p = \{\, \{a\}, \{a, b\}\,\} = \{\, \{a\}\,\}$.
>
> But then (c, d) also has only one member, because $(c, d) = p$. And this implies that $d = c$.
>
> So $d = c$ and $b = a$, and we already know that $c = a$.
>
> Hence $\boxed{d = b}$.
>
> Now assume that $b \neq a$.
>
> Then the sets $\{a\}$ and $\{a, b\}$ are different, because $b \in \{a, b\}$ but $b \notin \{a\}$.
>
> So p has two different members.
>
> And b is the only object that belongs to one of the members of p but does not belong to both.

And, similarly, d is the only object that belongs to one of the members of p but does nto belong to both.

So $\boxed{d = b}$.

We have proved that $d = b$ in both cases, when $b = a$ and when $b \neq a$.
So $\boxed{\boxed{d = b}}$.

So we have proved that $\boxed{\boxed{\boxed{c = a \wedge d = b}}}$. **Q.E.D.**

21.4.2 The Cartesian product of two sets

Definition 53. *Let A, B be sets. The Cartesian product of A and B is the set $A \times B$ given by*

$$A \times B = \left\{ u : (\exists a)(\exists b)(a \in A \wedge b \in B \wedge u = (a, b)) \right\}.$$

In other words, $A \times B$ (read as "A times B") is the set of all objects u such that u is an ordered pair (a, b), with $a \in A$ and $b \in B$.

Or, more succintly and elegantly, $A \times B$ ***is the set of all ordered pairs*** (a, b) ***for which*** $a \in A$ ***and*** $b \in B$.

Example 112.

- Let $A == \{1, 2, 3\}$ and $B = \{2, 3, 4, 5\}$. Then

$$A \times B = \left\{ (1,2), (1,3), (1,4), (1,5), (2,2), (2,3), (2,4), (2,5), (3,2), (3,3), (3,4), (3,5) \right\}.$$

  Notice that A is a finite set with 3 members, B is a finite set with 4 members, and $A \times B$ is a finite set with 12 members. ***This is not a coincidence. We will prove later that: if A, B are finite sets, A has m members, and B has n members, then $A \times B$ is a finite set and $A \times B$ has mn members.***

- Let $A = \mathbb{R}$, $B = \mathbb{R}$. Then $A \times B$ is $\mathbb{R} \times \mathbb{R}$, that is, the set of all ordered pairs (x, y) such that x and y are real numbers. ***This is the "x-y plane" of plane Euclidean geometry. The members of $\mathbb{R} \times \mathbb{R}$ are the "points" of plane geometry.***

- Let

$$\begin{aligned} A &= \{x \in \mathbb{R} : 0 < x < 1\}, \\ B &= \{x \in \mathbb{R} : 1 < x < 3\}. \end{aligned}$$

  Then A is the open interval $(0,1)$ (***not to be confused with the ordered pair*** $(0,1)$***!***) and B is the open interval $(1,3)$ (***not to be confused with the ordered pair*** $(1,3)$***!***). In this case, $A \times B$, that is, $(0,1) \times (1,3)$, is the set of all pairs (x,y) of real numbers such that $0 < xx < 1$ and $1 < y < 3$. In other words, $(0,1) \times (1,3)$ ***is the rectangle*** R ***characterized by the inequalities***

$$0 < x < 1 \qquad \text{and} \qquad 1 < y < 3\,.$$

21.5 Important facts about the set operations

So far, we have defined:

- One very special set (the empty set),
- One binary predicate (i.e., relation), about sets, namely, the predicate "is a subset of".
- Five binary operations on sets (union, intersection, difference, symmetric difference, and Cartesian product),
- One unary operation on sets (the power set).

By combining these nine things we can produce an enormous number of possible facts, some of which might be true, while others are not true. It would be pointless for me to give you a complete list and prove them all, because there are so many of them, and they are all so easy to prove (if true) or to disprove (if false).

And it would be pointless for you to memorize them all, because the list is so long. On the other hand, if you understand what yoiu are doing, you ought to be able, in each case, to figure out if the statement is true or false, and how to prove it (if it is true) or disprove it (if it is false).

So what I suggest is this: ***read carefully the list of facts, and pick a few of them and prove them or disprove them. Keep in mind that any of these facts could show up as a question in the exams.***

And here is the list:

1. If A is a set, then $\emptyset \subseteq A$. (True)

2. If A is a set, then $\emptyset \in A$. (False)

3. If A is a set, then $A \cup \emptyset = A$. (True)
NOTE: If you think that $\emptyset$ is like the number 0, and the operation "$\cup$" is like addition, then this statement is analogous to the statement that $x + 0 = x$ for every real number x.

4. If A is a set, then $A \cup \emptyset = \emptyset$. (False)

5. If A is a set, then $A \cap \emptyset = A$. (False)

6. If A is a set, then $A \cap \emptyset = \emptyset$. (True)
NOTE: If you think that $\emptyset$ is like the number 0, and the operation "$\cap$" is like multiplication, then this statement is analogous to the statement that $x.0 = 0$ for every real number x.

7. If A is a set, then $A \subseteq A$. (True)

8. If A, B are sets, then $A = B$ if and only if $A \subseteq B \wedge B \subseteq A$. (True)
NOTE: This gives an another way to prove that two sets are equal: to prove that $A = B$, you prove that $A \subseteq B$ and that $B \subseteq A$.

9. If A is a set, then $A \cup A = A$. (True)

10. If A is a set, then $A \cap A = A$. (True)

11. If A, B are sets, then $A \subseteq A \cup B$. (True)

12. If A, B are sets, then $A \subseteq A \cap B$. (False)

13. If A, B are sets, then $A \cup B \subseteq A$. (False)

14. If A, B are sets, then $A \cap B \subseteq A$. (True)

15. If A is a set, then $A \subseteq A$. (True)
NOTE: This aays that the binary relation "$\subseteq$" is reflexive.

16. If A, B are sets, $A \subseteq B$, and $B \subseteq A$, then $A = B$. (True)
NOTE: This aays that the binary relation "$\subseteq$" is antisymmetric.

17. If A, B, C are sets, $A \subseteq B$, and $B \subseteq C$, then $A \subseteq C$. (True)
NOTE: This aays that the binary relation "$\subseteq$" is transitive.

18. If A, B, C are sets, $A \subseteq B$, $B \subseteq C$, and $C \subseteq A$, then $A = B = C$. (True)

19. If A, B are sets, then $A \subseteq B$ if and only if $A \cup B = B$. (True)

20. If A, B are sets, then $A \subseteq B$ if and only if $A \cup B = A$. (False)

21. If A, B are sets, then $A \subseteq B$ if and only if $A \cap B = A$. (True)

22. If A, B are sets, then $A \subseteq B$ if and only if $A \cap B = B$. (False)

23. If A, B are sets, then $A \cup B = B \cup A$. (True)
NOTE: This is the ***commutative law of the union operation****.*

24. If A, B are sets, then $A \cap B = B \cap A$. (True)
NOTE: This is the ***commutative law of the intersection operation****.*

25. If A, B, C are sets, then $A \cup (B \cup C) = (A \cup B) \cup C$. (True)
NOTE: This is the ***associative law of the union operation****.*

26. If A, B, C are sets, then $A \cap (B \cap C) = (A \cap B) \cap C$. (True)
NOTE: This is the ***associative law of the intersection operation****.*

27. If A, B, C are sets, and $A \subseteq B$, then $A \cup C \subseteq B \cup C$. (True)

28. If A, B, C are sets, and $A \subseteq B$, then $A \cap C \subseteq B \cap C$. (True)

29. If A, B, C are sets, then $(A \cup B) \cap C = A \cup (B \cap C)$. (False)

30. If A, B, C are sets, then $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.(True)
NOTE: This is the ***distributive law of union with respect to intersection****.*

31. If A, B, C are sets, then $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.(True)
NOTE: This is the ***distributive law of intersection with respect to union****.*

    IMPORTANT NOTE: We have seen that union and intersection are in some ways like addition and multiplication: they obey commutative and associative laws. and also $A \cap \emptyset = \emptyset$ (which is analogous to $x \cdot 0 = 0$) and $A \cup \emptyset = A$ (which is analgous to $x + 0 = x$). But ***the analogy should not be pushed too far:***

    - *there is a distributive law of union with respect to intersection (i.e., $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$),*
    - *and there is also a distributive law of intersection with respect to union (i.e., $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$),*

- *but this is totally unlike what happens for addition and multiplication, because*
- *there is a distributive law of multiplication with respect to addition (i.e., $x \cdot (y + z) = x \cdot y + x \cdot z$)*
- *but there is no distributive law of addition with respect to multiplication (i.e., it is not true that $x+(y\cdot z) = (x+y)\cdot(x+z)$ since, for example, if we take $x = 1$, $y = 2$, $z = 3$, then $x + (y \cdot z) = 7$ and $(x + y) \cdot (x + z) = 12$).*

32. If A, B are sets, then $(A - B) \cup B = A$. (False)

33. If A, B are sets, then $(A - B) \cup B \subseteq A$. (False)

34. If A, B are sets, then $A \subseteq (A - B) \cup B$. (True)

35. If A, B, C are sets, then $A \cup (B - C) = (A \cup B) - (A \cup C)$. (False)

36. If A, B, C are sets, then $A \cap (B - C) = (A \cap B) - (A \cap C)$. (False)

    When we fix a "universe" U, then the complement of a subset A of U is defined to be the set $U - A$. The complement of A is denoted by "A^c".

37. If A, U are sets, and $A \subseteq U$, then $(A^c)^c = A$. (True)

38. If A, U are sets, and $A \subseteq U$, then $A \cup A^c = U$. (True)

39. If A, U are sets, and $A \subseteq U$, then $A \cap A^c == \emptyset$. (True)

40. If A, B, U are sets, $A \subseteq U$, and $B \subseteq U$, then

$$(A \cup B)^c = A^c \cap B^c \,. \tag{21.472}$$

    (This is true.)

41. If A, B, U are sets, $A \subseteq U$, and $B \subseteq U$, then

$$(A \cap B)^c = A^c \cup B^c \,. \tag{21.473}$$

    (This is true.)

    *NOTE: Equations (21.472) and (21.473) are the famous **De Morgan laws**. They say that*

- ***the complement of the union of two sets is the intersection of the complements of the sets,***

and

- ***the complement of the intersection of two sets is the union of the complements of the sets.***

I strongly recommend that you read the article on "De Morgab laws" in Wikipedia.

42. If A, B, U are sets, $A \subseteq U$, and $B \subseteq U$, then $A - B = A \cap B^c$. (True)

43. If A, B are sets, then $A - B = B - A$. (False)

44. If A, B, C are sets, then $A - (B - C) = (A - B) - C$. (False)

45. If A, B are sets, then $A \Delta B = (A \cup B) - (A \cap B)$. (True)

46. If A, B are sets, then $A \Delta B = B \Delta A$. (True)

47. If A, B, C are sets, then $A \Delta (B \Delta C) = (A \Delta B) \Delta C$. (True)

48. If A, B are sets, then $A \times B = B \times A$. (False)

49. If A, B, C, D are sets, then

$$(A \times B) \cap (C \times D) = (A \cap C) \times (B \cap D)\,.$$

(This is true.)

50. If A, B, C, D are sets, then

$$(A \times B) \cup (C \times D) = (A \cup C) \times (B \cup D)\,.$$

(This is false.)

51. If A is a set, then $A \in \mathcal{P}(A)$. (True)

52. If A is a set, then $A \subseteq \mathcal{P}(A)$. (False)

53. If A is a set, then $\emptyset \in \mathcal{P}(A)$. (True)

54. If A is a set, then $\emptyset \subseteq \mathcal{P}(A)$. (True)

55. If A, B are sets, then $A \subseteq B$ if and only if $\mathcal{P}(A) \subseteq \mathcal{P}(B)$. (True)

56. If A, B are sets, then $A = B$ if and only if $\mathcal{P}(A) = \mathcal{P}(B)$. (True)

57. If A, B are sets, then $\mathcal{P}(A \times B) = \mathcal{P}(A) \times \mathcal{P}(B)$. (False)

58. If A is a set, then $\emptyset \times A = \emptyset$ and $A \times \emptyset = \emptyset$. (True)

59. If A, B are sets, and $A \times B = B \times A$, then $A = B$. (False)

60. If A, B are nonempty sets, and $A \times B = B \times A$, then $A = B$. (True)

21.6 Some examples of proofs about sets

Let me give you the proofs of some of the results in the long list of the previous section.

21.6.1 Proof of one of the distributive laws

Theorem 88. *If A, B, C are sets, then*

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)\,. \tag{21.474}$$

Proof. To prove that the sets $A \cup (B \cap C)$ and $(A \cup B) \cap (A \cup C)$ are equal, we prove that they have the same members, that is, we prove that

$$(\forall x)\Big(x \in A \cup (B \cap C) \iff x \in (A \cup B) \cap (A \cup C)\Big)\,. \tag{21.475}$$

Sentence (21.475) is a universal sentence, of the form $(\forall x)P(x)$. So, in order to prove it, we let x be an arbitrary object and prove $P(x)$.

Let x be arbitrary.

We want to prove

$$x \in A \cup (B \cap C) \iff x \in (A \cup B) \cap (A \cup C)\,. \tag{21.476}$$

(1) The sentence "$x \in A \cup (B \cap C)$" is equivalent to "$x \in A \vee x \in B \cap C$". (Reason: if X, Y are sets, then the criterion for membership in $X \cup Y$ is "$x \in X \vee x \in Y$".)

(2) And "$x \in B \cap C$" is equivalent to "$x \in B \wedge x \in C$". (Reason: if X, Y are sets, then the criterion for membership in $X \cap Y$ is "$x \in X \wedge x \in Y$".)

(3) Hence "$x \in A \cup (B \cap C)$" is equivalent to "$x \in A \vee (x \in B \wedge x \in C)$".

(4) Also, "$x \in (A \cup B) \cap (A \cup C)$" is equivalent to "$x \in A \cup B \wedge x \in A \cup C$".

And

- "$x \in A \cup B$" is equivalent to "$x \in A \vee x \in B$".
- "$x \in A \cup C$" is equivalent to "$x \in A \vee x \in C$".

(5) So "$x \in (A \cup B) \cap (A \cup C)$" is equivalent to "$(x \in A \vee x \in B) \wedge (x \in A \vee x \in C)$".

It follows from (3) and (6) that "$x \in A \cup (B \cap C) \iff x \in (A \cup B) \cap (A \cup C)$", the sentence that we have to prove, is equivalent to

$$x \in A \vee (x \in B \wedge x \in C) \iff \Big((x \in A \vee x \in B) \wedge (x \in A \vee x \in C) \Big) . \tag{21.477}$$

The sentence (21.477) is of the form

$$P \vee (Q \wedge R) \iff \Big(P \vee Q) \wedge (P \vee R) \Big) , \tag{21.478}$$

where P stands for "$x \in A$", Q stands for "$x \in B$", and R stands for "$x \in C$".

We now prove that (21.478) is true.

Sentence (21.478) is a biconditional, of the form $\mathcal{L} \iff \mathcal{M}$. And a biconditional $\mathcal{L} \iff \mathcal{M}$ is true if and only if $\mathcal{L}$ and $\mathcal{M}$ have the same truth value, i.e., are both true or both false. So we are going to prove that $\mathcal{M}$ is true if $\mathcal{L}$ is true and $\mathcal{M}$ is false if $\mathcal{L}$ is false.

> Suppose that $P \vee (Q \wedge R)$ is true.
>
> Then either P is true or $Q \wedge R$ is true.
>
> > Suppose P is true.
> > Then both $P \vee Q$ and $P \vee R$ are true.
> > So $(P \vee Q) \wedge (P \vee R)$ is true.
> > Now suppose that $Q \wedge R$ is true.
> > Then both Q and R are true.
> > So $P \vee Q$ and $P \vee R$ are true.
> > And then $(P \vee Q) \wedge (P \vee R)$ is true.
>
> So $(P \vee Q) \wedge (P \vee R)$ is true in both cases, and then $(P \vee Q) \wedge (P \vee R)$ is true.

This proves that $(P \vee Q) \wedge (P \vee R)$ is true if $P \vee (Q \wedge R)$ is true.

Now suppose that $P \vee (Q \wedge R)$ is false.

Then both P and $Q \wedge R$ are false.

Since $Q \wedge R$ is false, either Q is false or R is false.

Suppose Q is false.

Since P is false, $P \vee Q$ is false, because both P and Q are false.

Hence the conjunction $(P \vee Q) \wedge (P \vee R)$ is false.

Now suppose R is false.

Since P is false, $P \vee R$ is false, because both P and R are false.

Hence the conjunction $(P \vee Q) \wedge (P \vee R)$ is false.

So $(P \vee Q) \wedge (P \vee R)$ is false in both cases, and then $(P \vee Q) \wedge (P \vee R)$ is false.

This proves that $(P \vee Q) \wedge (P \vee R)$ is false if $P \vee (Q \wedge R)$ is false.

So we have proved that (21.477) is true, and this completes our proof, **Q.E.D.**

Problem 108. ***Prove*** the other distributive law: If A, B, C are sets, then

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)\,. \tag{21.479}$$

21.6.2 Proofs of the De Morgan laws

As we explained before, the De Morgan laws are the following two statements.

Theorem 89. *Let U be a set, and let A, B be subsets of U. Then*

$$(A \cup B)^c \;=\; A^c \cap B^c\,,$$

and

Theorem 90. *Let U be a set, and let A, B be subsets of U. Then*

$$\begin{aligned}(A \cup B)^c &= A^c \cap B^c\,,\\ (A \cap B)^c &= A^c \cap B^c\,.\end{aligned}$$

I will give you a proof from first principles[116] of the first theorem, and then I will give you a short proof of the other using the first one, and ask you to give a proof from first principles of the second theorem.

Proof. We want to prove that

$$DeMorgan (\forall x \in U) \Big(x \in (A \cup B)^c \iff x \in A^c \cap B^c \Big) . \tag{21.480}$$

The sentence we want to prove is a universal sentence, of the form $(\forall x) P(x)$. So in order to prove it we let x be an arbitrary object and prove $P(x)$.

Let x be an arbitrary member of U.

We want to prove that

$$x \in (A \cup B)^c \iff x \in A^c \cap B^c . \tag{21.481}$$

But, for $x \in U$, "$x \in (A \cup B)^c$ is equivalent to "$x \notin A \cup B$", i.e., to "$\sim x \in A \cup B$".

And "$x \in A \cup B$" is equiva;ent to "$x \in A \vee x \in B$".

So "$x \notin A \cup B$" is equivalent to "$\sim (x \in A \vee x \in B)$".

Therefore "$x \in (A \cup B)^c$" is equivalent to "$\sim (x \in A \vee x \in B)$".

On the other hand, "$x \in A^c \cap B^c$" is equiva;ent to "$x \in A^c \wedge x \in B^c$".

And the sentences "$x \in A^c$", "$x \in B^c$" are equivalent to "$\sim x \in A$" and "$\sim x \in B$".

So "$x \in A^c \cap B^c$" is equivalent to "$(\sim x \in A) \wedge (\sim x \in B)$".

Hence (21.481) is equivalent to

$$\Big(\sim (x \in A \vee x \in B) \Big) \iff \Big((\sim x \in A) \wedge (\sim x \in B) \Big) . \tag{21.482}$$

If we use P to stand for "$x \in A$", and Q to stand for "$x \in B$", then (21.482) is the sentence

$$\Big(\sim (P \vee Q) \Big) \iff \Big((\sim P) \wedge (\sim Q) \Big) . \tag{21.483}$$

[116] A bf*proof from first principles* is a proof in which you do not use any intermediate results proved before. For example, after we proved that $2 + 2 = 4$ from first principles we proved that $2 \times 2 = 4$ using the result that $2 + 2 = 4$. That was ***not*** a proof from first principles. In a proof from first principles, you would just have used the basic facts and the definitions, and no theorem proved before.

The biconditional sentence (21.483) is of the form $\mathcal{L} \iff \mathcal{M}$. And a biconditional $\mathcal{L} \iff \mathcal{M}$ is true if and only if $\mathcal{L}$ and $\mathcal{M}$ have the same truth value, i.e., are both true or both false. So we are going to prove that $\mathcal{M}$ is true if $\mathcal{L}$ is true and $\mathcal{M}$ is false if $\mathcal{L}$ is false.

Proof that if $\sim (P \vee Q)$ *is true then* $(\sim P) \wedge (\sim Q)$ *is true.*

Suppose that $\boxed{\sim (P \vee Q) \text{ is true}}$.

Then $P \vee Q$ is false.

So both P and Q are false.

Hence $\sim P$ and $\sim Q$ are true.

So the conjunction $\boxed{(\sim P) \wedge (\sim Q) \text{ is true}}$.

Proof that if $\sim (P \vee Q)$ *is false then* $(\sim P) \wedge (\sim Q)$ *is false.*

Suppose that $\boxed{\sim (P \vee Q) \text{ is false}}$.

Then $P \vee Q$ is true.

So either P is true or Q is true.

Suppose that P is true.

Then $\sim P$ is false.

So the conjunction $(\sim P) \wedge (\sim Q)$ is false.

Now suppose that Q is true.

Then $\sim Q$ is false.

So the conjunction $(\sim P) \wedge (\sim Q)$ is false.

We have shown that $(\sim P) \wedge (\sim Q)$ is false in both cases, when P is true and when Q is true.

Hence $\boxed{(\sim P) \wedge (\sim Q) \text{ is false}}$.

So we have proved (21.481) for an arbitrary member x of U, and we can go to

$$(\forall x \in U)\Big(x \in (A \cup B)^c \iff x \in A^c \cap B^c\Big)\,. \tag{21.484}$$

And (21.484) says that the sets $(A \cup B)^c$ and $(A \cap B)^c$ have rhe same members, so the sets are equal, that ism

$$(A \cup B)^c = A^c \cap B^c\,. \tag{21.485}$$

This is exactly what we wanted to prove. **Q.E.D.**

Now let us give a simple proof of Theorem 90 using Theorem 89.
Proof. We want to prove that $(A \cap B)^c = A^c \cup B^c$.

Theorem 89 says that, if X, Y are any subsets of U, then

$$(X \cup Y)^c = X^c \cap Y^c \,. \tag{21.486}$$

Apply this with $X = A^c$ and $B = Y^c$. We get

$$(A^c \cup B^c)^c = (A^c)^c \cap (B^c)^c \,. \tag{21.487}$$

But $(A^c)^c = A$, and $(B^c)^c = B$. So

$$(A^c \cup B^c)^c = A \cap B \,. \tag{21.488}$$

Now take the complement of both sides. We get

$$\Big((A^c \cup B^c)^c\Big)^c = (A \cap B)^c \,. \tag{21.489}$$

But $(X^c)^c = X$ for every subset X of U. Therefore

$$\Big((A^c \cup B^c)^c\Big)^c = A^c \cup B^c \tag{21.490}$$

Combining (21.489) and (21.490), we get

$$A^c \cup B^c c = (A \cap B)^c \,, \tag{21.491}$$

which is the formula we were trying to prove. **Q.E.D.**

Problem 109. Write a proof from first principles of Theorem 90. ***I strongly recommend that you use the same style as in the proof of Theorem 89. The proof of Theorem 89 is really very simple, and almost mechanical. It looks long because it was written on purpose to show you a proof written in a very precise, very detailed way, displaying the use of the rules of logic. Usually one does not write p;roofs like that, but I would like you to do it at least once, to show that you can do it.*** □

21.6.3 A proof involving the symmetric difference

Let us prove Fact 45 from our list. Recall that the ***symmetric difference*** of two sets A, B is the set $A\Delta B$ given by

$$A\Delta B = (A - B) \cup (B - A) \,.$$

In the proof, we are going to use the following facts, that are valid for arbitrary subsets X, Y, Z of a set U:

- $X - Y = X \cap Y^c$,
- $X \cup X^c = U$ and $X \cap X^c = \emptyset$,
- $X \cap U = X$ and $X \cap \emptyset = \emptyset$.
- $X \cup U = U$ and $X \cup \emptyset = X$.
- The commutative laws

$$\begin{aligned} X \cup Y &= Y \cup X\,, \\ X \cap Y &= Y \cap X\,, \end{aligned}$$

- The associative laws

$$\begin{aligned} X \cup (Y \cup Z) &= (X \cup Y) \cup Z\,, \\ X \cap (Y \cap Z) &= (X \cap Y) \cap Z\,, \end{aligned}$$

- The distributive laws

$$\begin{aligned} X \cup (Y \cap Z) &= (X \cup Y) \cap (X \cup Z)\,, \\ X \cap (Y \cup Z) &= (X \cap Y) \cup (X \cap Z)\,, \end{aligned}$$

- The De Morgan laws

$$\begin{aligned} X^c \cup Y^c &= (X \cap Y)^c\,, \\ X^c \cap Y^c &= (X \cup Y)^c\,. \end{aligned}$$

Theorem 91. *If A, B are sets, then $A \Delta B = (A \cup B) - (A \cap B)$.*

Proof. Choose as universe any set U such that $A \subseteq U$ and $B \subseteq U$. (For example, we could choose U to be $A \cup B$.)

Then

$$\begin{aligned}
A\Delta B &= (A-B)\cup(B-A) && (21.492)\\
&= (A\cap B^c)\cup(B\cap A^c) && (21.493)\\
&= \Big((A\cap B^c)\cup B\Big)\cap\Big((A\cap B^c)\cup A^c\Big) && (21.494)\\
&= \Big(B\cup(A\cap B^c)\Big)\cap\Big(A^c\cup(A\cap B^c)\Big) && (21.495)\\
&= \Big((B\cup A)\cap(B\cup B^c)\Big)\cap\Big((A^c\cup A)\cap(A^c\cup B^c)\Big) && (21.496)\\
&= \Big((B\cup A)\cap U\Big)\cap\Big(U\cap(A^c\cup B^c)\Big) && (21.497)\\
&= (B\cup A)\cap(A^c\cup B^c) && (21.498)\\
&= (A\cup B)\cap(A^c\cup B^c) && (21.499)\\
&= (A\cup B)\cap(A\cap B)^c\,. && (21.500)
\end{aligned}$$

So $A\Delta B = (A\cup B)-(A\cap B)$. **Q.E.D.**

Problem 110. Write the justifications of each of the nine steps (21.492), (21.493), (21.494), (21.495), (21.496), (21.497), (21.498), (21.499), (21.500) of the proof of Theorem 91. □

Problem 111. ***Prove or disprove*** each of the following distributive laws

1. *The distributive law of intersection with respect to symmetric difference.* If A, B, C are sets, then

$$A\cap(B\Delta C) = (A\cap B)\Delta(A\cap C)\,. \qquad (21.501)$$

2. *The distributive law of union with respect to symmetric difference.* If A, B, C are sets, then

$$A\cup(B\Delta C) = (A\cup B)\Delta(A\cup C)\,. \qquad (21.502)$$

Appendix: a lemma on rearranging lists of numbers

First of all, let us introduce the notion of "equivalent lists".

Definition 54. *Let* $\mathbf{p} = (p_j)_{j=1}^n$ *and* $\mathbf{q} = (q_j)_{j=1}^m$ *be finite lists. We say that* $\mathbf{p}$ *and* $\mathbf{q}$ *are* equivalent *(or that* $\mathbf{p}$ *is a* rearrangement *of* $\mathbf{q}$*, or that* $\mathbf{q}$ *is a* rearrangement *of* $\mathbf{p}$*) if*

1. $m = n$,
2. *the sets*

$$\begin{aligned} \mathrm{Set}(\mathbf{p}) &= \{x : (\exists j \in \mathbb{N}_m) p_j = x\}, \\ \mathrm{Set}(\mathbf{q}) &= \{x : (\exists j \in \mathbb{N}_m) q_j = x\}, \end{aligned}$$

   are equal,

3. *every member of* $\mathrm{Set}(\mathbf{p})$ *(i.e., of* $\mathrm{Set}(\mathbf{q})$*) occurs the same number of times as an entry of* $\mathbf{p}$ *as it does as an entry of* $\mathbf{q}$. □

We will write

$$\mathbf{p} \equiv \mathbf{q}$$

to indicate that $\mathbf{p}$ is a rearrangement of $\mathbf{q}$.

(II) **Lemma 2**. *Let* $\mathbf{p} = (p_j)_{j=1}^n$ *be a finite list of real numbers. Then there exists a list* $\mathbf{q} = (q_j)_{j=1}^n$ *such that*

1. $\mathbf{q} \equiv \mathbf{p}$,
2. $\mathbf{q}$ *is ordered,*
3. $\sum_{j=1}^n p_j = \sum_{j=1}^n q_j$,
4. $\prod_{j=1}^n p_j = \prod_{j=1}^n q_j$.

Proof. We do a proof by induction.

Let $P(n)$ be the statement

> For every list $\mathbf{p} = (p_j)_{j=1}^n$ of length n consisting of real numbers there exists an ordered list $\mathbf{q} = (q_j)_{j=1}^n$ that is equivalent to $\mathbf{p}$ and satisfies $\sum_{j=1}^n p_j = \sum_{j=1}^n q_j$ and $\prod_{j=1}^n p_j = \prod_{j=1}^n q_j$.

We prove that $(\forall n \in \mathbb{N}) P(n)$ by induction on n.

The base case. $P(1)$ is obviously true, because if $\mathbf{p} = (p_j)_{j=1}^{1}$ is a list having just one entry, then of course $\mathbf{p}$ is ordered, so we can take $\mathbf{q}$ to be $\mathbf{p}$, and then $\mathbf{q}$ is an ordered list and is equivalent to $\mathbf{p}$.

The inductive step. We want to prove $(\forall n \in \mathbb{N})(P(n) \implies P(n+1))$.

Let $n \in \mathbb{N}$ be arbitrary.

Assume that $P(n)$ is true.

We want to prove $P(n+1)$.

Statement $P(n+1)$ says

$$(\forall \mathbf{p})\Bigg(\mathbf{p} = (p_j)_{j=1}^{n+1} \text{ is a list of real numbers} \implies$$
$$(\exists \mathbf{q})\Big(\mathbf{q} = (q_j)_{j=1}^{n+1} \text{ is a list of length } n+1 \wedge \mathbf{q} \text{ is ordered} \wedge$$
$$\mathbf{q} \equiv \mathbf{p} \wedge \sum_{j=1}^{n+1} p_j = \sum_{j=1}^{n+1} q_j \wedge \prod_{j=1}^{n+1} p_j = \prod_{j=1}^{n+1} q_j\Big)\Bigg).$$

To prove $P(n+1)$ we must take an arbitrary $\mathbf{p}$, assume that $\mathbf{p}$ is a list of real numbers of lenght $n+1$, and prove that there exists an ordered list $\mathbf{q}$ that is equivalent to $\mathbf{p}$ and satisfies the conditions on the sum and the product.

Let $\mathbf{p}$ be an arbitrary list of real numbers of length $n+1$.

Let $\mathbf{p} = (p_j)_{j=1}^{n+1}$.

Let j_* be an index belonging to $\mathbb{N}_{n+1}$ such that p_{j_*} has the maximum possible value of all the p_j. (That is, precisely[117], $j_* \in \mathbb{N}_{n+1}$ and $p_{j_*} = \operatorname{Max} \mathbf{p}$.)

Let $\mathbf{p}'$ be the list of length n obtained from $\mathbf{p}$ by removing the j_*-th entry. (Precisely, let $\mathbf{p}' = (p'_j)_{j=1}^{n}$ be the list defined by $p'_j = p_j$ for $j < j_*$, and $p'_j = p_{j+1}$ for $j_* \le j \le n$.)

Then $\mathbf{p}'$ is a list of primes of length n.

Since we are assuming that $P(n)$ holds, there exists an ordered list $\mathbf{q}' = (q'_j)_{j=1}^{n}$ such that $\mathbf{q}' \equiv \mathbf{p}'$, $\sum_{j=1}^{n} q'_j = \sum_{j=1}^{n} p'_j$, and $\prod_{j=1}^{n} q'_j = \prod_{j=1}^{n} p'_j$.

[117] The existence of such a j_* is a consequence of Theorem 82. This theorem says that every finite list of real numbers has a lagest entry, which is completely obvious, but can also be proved rigorously if anyone so desires.

Let $\mathbf{p}''$ be the list of length $n+1$ obtained from $\mathbf{p}'$ by adding p_{j_*} as the $n+1$-th entry. (Precisely, $\mathbf{p}'' = (p''_j)_{j=1}^{n+1}$, where $p''_j = p'_j$ for $j \in \mathbb{N}$, and $p''_{n+1} = p_{j_*}$.)
Let $\mathbf{q}''$ be the list of length $n+1$ obtained from $\mathbf{q}'$ by adding p_{j_*} as the $n+1$-th entry. (Precisely, $\mathbf{q}'' = (q''_j)_{j=1}^{n+1}$, where $q''_j = q'_j$ for $j \in \mathbb{N}$, and $q''_{n+1} = p_{j_*}$.)
Since $\mathbf{q}' \equiv \mathbf{p}'$ and the lists $\mathbf{q}''$, $\mathbf{p}''$ are obtained from $\mathbf{q}'$ and $\mathbf{p}'$ by adding the same entry p_{j_*} at the end, it is clear that $\mathbf{q}'' \equiv \mathbf{p}''$.
Since $\mathbf{p}''$ is obtained from $\mathbf{p}$ by interchanging two entries (by moving p_{j_*} from the j_*-th position to the $n+1$-th position), it is clear that $\mathbf{p}'' \equiv \mathbf{p}$.
So $\mathbf{q}'' \equiv \mathbf{p}$.
Furthermore, $\mathbf{q}''$ is ordered. (Reason: $\mathbf{q}'$ is ordered, so the first n entries of $\mathbf{q}''$ satisfy $q''_1 \le q''_2 \le \cdots \le q''_n$. In addition, for some $j \in \mathbb{N}_{n+1}$, $q''_n = p_j \le p_{j_*} = q''_{n+1}$.)
Finally,

$$\begin{aligned}
\sum_{j=1}^{n+1} q''_j &= (\sum_{j=1}^{n} q''_j) + q''_{n+1} = (\sum_{j=1}^{n} q'_j) + p_{j_*} = (\sum_{j=1}^{n} p'_j) + p_{j_*} \\
&= (\sum_{j=1}^{j_*-1} p'_j + \sum_{j=j_*}^{n} p'_j) + p_{j_*} = (\sum_{j=1}^{j_*-1} p_j + \sum_{j=j_*}^{n} p_{j+1}) + p_{j_*} \\
&= (\sum_{j=1}^{j_*-1} p_j + \sum_{j=j_*+1}^{n+1} p_j) + p_{j_*} \\
&= (\sum_{j=1}^{j_*-1} p_j) + p_{j_*} + (\sum_{j=j_*+1}^{n+1} p_j) \\
&= \sum_{j=1}^{n+1} p_j \,,
\end{aligned}$$

so

$$\sum_{j=1}^{n+1} q''_j = \sum_{j=1}^{n+1} p_j \,.$$

$*$ A similar argument shows that

$$\prod_{j=1}^{n+1} q''_j = \prod_{j=1}^{n+1} p_j \,.$$

So, if we take $\mathbf{q}$ to be $\mathbf{q}''$, we have shown that $\mathbf{q}$ satisfies all the conditions that appear in statement $P(n+1)$.

This completes the proof of $P(n+1)$, assuming $P(n)$.

Hence $P(n) \implies P(n+1)$.

- So $(\forall n \in \mathbb{N})(P(n) \implies P(n+1))$.

This completes the inductive step, and the proof of our lemma. **Q.E.D.**

22 Relations and functions

22.1 The definition of "relation"

Definition 55. *A relation is a set of ordered pairs.* □

That is:

- a relation is a set R such that every member of R is an ordered pair,
- equivalently, a relation is a set R such that

$$(\forall x \in R)(\exists u)(\exists v)\, x = (u, v)\,. \tag{22.503}$$

You should picture a relation as set of arrows: for each u and each v, you draw an arrow from u to v to indicate that the pair (u, v) belongs to R.

Another way to think of a relation R is as some kind of device that takes in "inputs" and produces "outputs". The pairs belonging to R are the ***input-output pairs*** produced by R.

For example, for the the relation

$$R = \{\, (1,2), (1,3), (1,4), (2,1), (2,2), (3,1), (3,5), (4,5) \,\}\,, \tag{22.504}$$

- the input 1 will produce the outputs 2, 3, and 4,
- the input 2 will produce the outputs 1 and 2,
- the input 3 will produce the outputs 1 and 5,
- the input 4 will produce the output 5.

22.1.1 The domain and range of a relation

Definition 56. *If R is a relation, an input of R is an object x such that $(x, y) \in R$ for some y.* □

(That is, x is an input of R if $(\exists y)(x, y) \in R$.)

Definition 57. *If R is a relation, an output of R is an object y such that $(x, y) \in R$ for some x.* □

(That is, y is an output of R if $(\exists x)(x, y) \in R$.)

Definition 58. *If R is a relation and x is an input for R, an output of R for the input x is an object y such that $(x, y) \in R$.*

(That is, y is an output of R for x if $(x, y) \in R$.) □

Definition 59. *If R is a relation, the domain of R is the set of all inputs of R. We use* $\mathrm{Dom}(R)$ *to denote the domain of R. Therefore*

$$\mathrm{Dom}(R) = \{x : (\exists y)(x, y) \in R\}\,. \tag{22.505}$$

Definition 60. *If R is a relation, the range of R is the set of all outputs of R. We use* $\mathrm{Ran}(R)$ *to denote the range of R. Therefore*

$$\mathrm{Ran}(R) = \{y : (\exists x)(x, y) \in R\}\,. \tag{22.506}$$

22.2 Functions

22.2.1 The unique output property

Definition 61. *If R is a relation, and x is an input of R (i.e., $x \in \mathrm{Dom}(R)$), we say that x has the unique output property if there is only one output of R for x.*

That is, x has the unique output property if

$$(\forall y)(\forall z)\Big(\big((x, y) \in R \wedge (x, z) \in R\big) \Longrightarrow y = z\,.\Big) \tag{22.507}$$

Example 113. For the relation R given by (22.504), the input 4 has the unique output property. The inputs 1, 2 and 3 do not. □

22.2.2 The definition of "function"

Definition 62. *A function is a relation f such that every input of f has the unique output property.* □

That is, a set f is a function if

(i) f is a set of ordered pairs.

(ii) for all x, y, z, if $(x, y) \in f$ and $(x, z) \in f$ then $y = z$.

In purely formal language, f is a function if

(I) $(\forall u \in f)(\exists x)(\exists y)u = (x, y)$,

(II) $(\forall x)(\forall y)(\forall z)\Big(\big((x, y) \in f \wedge (x, z) \in f\big) \Longrightarrow y = z\Big)$.

22.2.3 The definition of "value" of a function at a member of its domain

Definition 63. *If f is a function and $x \in \mathrm{Dom}(f)$ (that is, x is an input of f), then the value of f at x is the object $f(x)$ such that $f(x)$ is the unique output of f for x. (Recall that the definition of function tells us that every input of f has the unique output property, so the output for x is indeed unique.)* □

22.2.4 When are two functions equal?

The following theorem is the analogue for functions of the theorem on equality of sets: two sets are the same set if they have the same members. Here, the result says: two functions are the same function if they have the same domain and, for each member x of this domain, have the same values at x.

Theorem 92. *Let f, g be functions. Then $f = g$ if and only if* $\mathrm{Dom}(f) = \mathrm{Dom}(g)$ *and $f(x) = g(x)$ for every $x \in \mathrm{Dom}(f)$.*

Proof. It is clear that of $f = g$ then $\mathrm{Dom}(f) = \mathrm{Dom}(g)$ and $f(x) = g(x)$ for every $x \in \mathrm{Dom}(f)$.

Now assume that $\mathrm{Dom}(f) = \mathrm{Dom}(g)$ and $f(x) = g(x)$ for every $x \in \mathrm{Dom}(f)$.

We want to prove that $f = g$. Since f and g are sets, it suffices to prove that $f \subseteq g$ and $g \subseteq f$. Both proofs are the same, so I will do only one of them.

Let us prove that $f \subseteq g$. Let u be an arbitrary member of f. Since f is a set of ordered pairs, u is an ordered pair, so we may pick x, y such that $u = (x, y)$. Since $(x, y) \in f$, x is an input of f, so $x \in \mathrm{Dom}(f)$, and y is an output of f for x, so $y = f(x)$. Since $\mathrm{Dom}(f) = \mathrm{Dom}(g)$, and $x \in \mathrm{Dom}(f)$, we see that $x \in \mathrm{Dom}(g)$. Since $g(x) = f(x)$, it follows that $(x, y) \in g$. So $u \in g$

So we have proved that every $u \in f$ is in g. Hence $f \subseteq g$. **Q.E.D.**

22.2.5 The definition of "function from a set to a set"

Definition 64. *If f, A, B are sets, we say that f is a function from A to B, and write*

$$f : A \to B\,,$$

if

1. *f is a function,*
2. *A is the domain of f,*
3. *The range of f is a subset of B.* □

In other words, "$f : A \to B$" means "f is a function, A is the domain of f, and $f(x) \in B$ for every $x \in A$".

22.2.6 Composition of functions

Definition 65. *If A, B, C are sets, $f : A \to B$ and $g : B \to C$, the composite function of f and g is the function $g \circ f : A \to C$ given by*

$$g \circ f(x) = g\big(f(x)\big) \qquad \text{for every} \quad x \in A\,.$$

The following theorem says that the operation of composition of functions satisfies the associative law in the sense that when both $h \circ (g \circ f)$ and $(h \circ g) \circ f$ are defined, they are equal

Theorem 93. *Let A, B, C, D be sets, and assume that $f : A \to B$, $g : B \to C$, and $h : C \to D$. Then*

$$h \circ (g \circ f) = (h \circ g) \circ f\,. \tag{22.508}$$

Proof. **YOU DO IT.**

Problem 112. ***Prove*** Theorem 93. □

22.2.7 The definition of "one-to-one function"

Definition 66. *A function f is <u>one-to-one</u> (or <u>injective</u>) if whenever two inputs are different, the values of f at those inputs are different as well.*

In other words: f is one-to-one if

$$(\forall u \in \mathrm{Dom}(f))(\forall v \in \mathrm{Dom}(f))(u \neq v \Longrightarrow f(u) \neq f(v))\,. \tag{22.509}$$

Equivalently, f is one-to-one if

$$(\forall u \in \mathrm{Dom}(f))(\forall v \in \mathrm{Dom}(f))\big(f(u) = f(v) \Longrightarrow u = v\big)\,. \tag{22.510}$$

22.2.8 The composite of two one-to-one functions

Theorem 94. *Let A, B, C be sets, and assume that $f : A \to B$, $g : B \to C$, and both f and g are one-to-one. Then $g \circ f$ is one-to-one.*

Proof. Let $h = g \circ f$. We want to prove that h is one-to-one.

For that purpose, we prove that if $x_1 \in A$, $x_2 \in A$, and $x_1 \neq x_2$, it follows that $h(x_1) \neq h(x_2)$.

Let $y_1 = f(x_1)$, $y_2 = f(x_2)$, Since f is one-to-one and $x_1 \neq x_2$, we can conclude that $y_1 \neq y_2$.

Then, since g is one-to-one and $y_1 \neq y_2$, we can conclude that $g(y_1) \neq g(y_2)$.

But $g(y_1) = g(f(x_1)) = h(x_1)$, and $g(y_2) = g(f(x_2)) = h(x_12)$. So $h(x_1) \neq h(x_2)$, as desired. **Q.E.D.**

22.2.9 The definition of "function onto a set"

Definition 67. *A function $f : A \to B$ is <u>onto B</u> if $B = \mathrm{Ran}(f)$.*

In other words, f is onto B if

$$(\forall b \in B)(\exists a \in A) f(a) = b\,. \tag{22.511}$$

Example 114. Let f be the "squaring a real number" function.

The domain $\mathrm{Dom}(f)$ is $\mathbb{R}$, the set of all real numbers. And, for $x \in \mathbb{R}$, the value $f(x)$ is given by

$$f(x) = x^2\,.$$

Then the range $\mathrm{Ran}(f)$ is $\mathbb{R}_+$, the set of all nonnegative real numbers. (Reason: every nonnegative real number has a square root.)

Then both statements "$f : \mathbb{R} \to \mathbb{R}$" and "$f : \mathbb{R} \to \mathbb{R}_+$" are true. And f is onto $\mathbb{R}_+$ but f is not onto $\mathbb{R}$. □

Remark 31. The previous example shows that the sentence "f is onto" is meaningless, in the same way as the the sentence "a is divisible" is meaningless.

There is no such thing as "being divisible". What makes sense is "being divisible by some number". "Divisible" is a 2-argument predicate: we say things like "a is divisible by b", and we do not say things like "a is divisible". Similarly, "is onto" is a 2-argument predicate: we say things like "f is onto B", and we do not say things like "f is onto". □

22.2.10 The composite of two onto functions

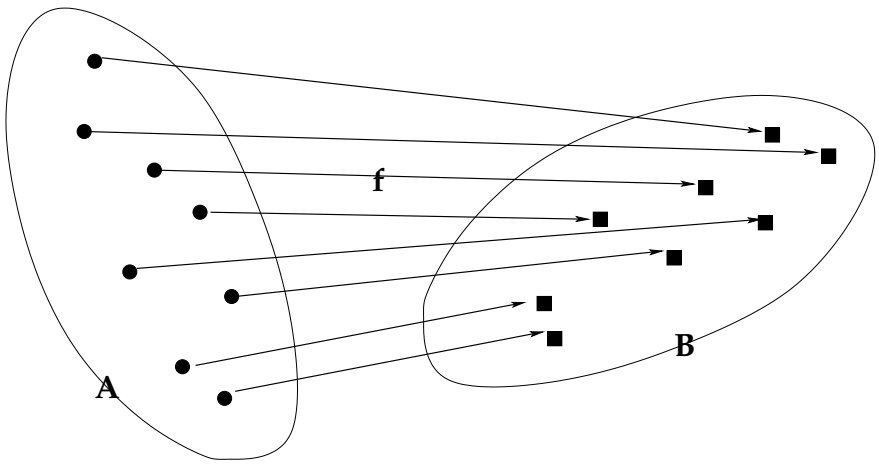
Theorem 95. *Let A, B, C be sets, and assume that $f : A \to B$, $g : B \to C$, f is onto B and g is onto C. Then $g \circ f$ is onto C.*

Proof. **YOU DO IT.**

Problem 113. ***Prove*** Theorem 95. □

22.3 The definition of "bijection"

Definition 68. *A function $f : A \to B$ is a <u>bijection from A to B</u> if it is one-to-one and onto B.* □



A bijection

We can tell that A (the set of circles) and B (the set of squares) have the same number of members because f (the set of arrows, i.e. of ordered pairs) is a bijection from A to B:

- the arrows send every circle to one and only one square (that is, f is a function),
- different circles are sent to different squares (that is, f is one-to-one),
- no square is left out (that is, f is onto).

Theorem 96. *The empty set is a bijection from the empty set to the empty set.*

Proof. **YOU PROVE THIS.**

Problem 114. ***Prove*** Theorem 96.

You have to prove several things:

1. that $\emptyset$ is a relation, i.e., a set of ordered pairs. (So you have to prove that every member of $\emptyset$ is an ordered pair.)
2. that $\emptyset$ is in fact a function. (So you have to prove that
$$(\forall x)(\forall y)(\forall z)\Big(\big((x,y)\in\emptyset\wedge(x,z)\in\emptyset\big)\Longrightarrow y=z\Big),$$
3. that the domain of $\emptyset$ is $\emptyset$,

4. that $\emptyset : \emptyset \to \emptyset$,
5. that $\emptyset$ is one-to-one,
6. that $\emptyset$ is onto $\emptyset$. □

22.3.1 The exchange lemma

The following theorem is rather simple, but very important. To understand what is says, think of of an example. Suppose you have several couples dancing in a large ballroom: every man is dancing with one and only one woman, every woman is danving with one and only one man. That means that we have in front of our eyes a bijection f from M to W, if M is the set of all the men that are dancing, and W is the set of all the women. The bijection $f : M \to W$ is defined as follows: for $m \in M$, $f(m)$ is the woman with whom m is dancing.

Theorem 97 then says that, if we are interested in a particular man a and a particular woman b, we can rearrange things so that a will be dancing with b. And the way we do that is as follows: if a is already dancing with b then we do not have to do anything. And if a is not dancing with b, that means that a is dancing with some other woman b', and b is dancing with some other man a'. So in this case we just have the two couples (a, b') (a', b) ***exchange partners***: we remove the pairs[118] (a, b'), (a', b) from f and put, instead, the pairs (a, b) and (a', b'), this gives rise ot a new bijection g. And $g(a)$ is b, as we wanted.

Theorem 97. *Suppose A, B are sets, $f : A \to B$, and f is a bijection from A to B. Then for every $a \in A$ and every $b \in B$ there exists a bijection g from A to B such that $g(a) = b$.*

Proof. **YOU PROVE THIS.**

Problem 115. ***Prove*** Theorem 97. Make sure you write a complete, detailed proof of the fact that the function g that you define is a bijection.

HINT: Read carefully the explanation before the statement of the theorem. All you have to do is repeat the same construction of g in general, for any sets A, B, and bijection f, without mentioning "men", "women", and "dancing". □

[118] Do not forget that f is a set of ordered pairs. The fact that a is dancing with b' and a' is dancing with b means that $b' = f(a)$ and $b = f(a')$, and this in turn means that the pairs (a, b') and (a', b) are members of f.

22.3.2 The composite of two bijections

Theorem 98. *Let A, B, C be sets, and assume that f is a bijection from A to B and g is a bijection from B to C. Then $g \circ f$ is a bijection from A to C.*

Proof. Let $h = g \circ f$. We want to prove that h is a bijection.

Since f and g are bijections, they are one-to-one. Hence h is one-to-one by Theorem 94.

Since f and g are bijections, f is onto B and g is onto C. Hence h is onto C by Theorem 95.

So h is one-to-one and onto C. Therefore h is a bijection from A to C. **Q.E.D**.

22.3.3 The identity function of a set

Definition 69. *If A is a set, the identity function of A, or identity map of A, is the function $1_A : A \to A$ such that*

$$1_A(x) = x \qquad \text{for every } x \in A\,. \tag{22.512}$$

The reason $!_A$ is called the "identity function" is that it behaves, with respect to the operation of function composition, very much like the number 1 behaves with respect to the operation of multiplication of numbers:

- Multiplying a number by 1 yields the same number: $x \cdot 1 = x$ for every number x.
- Composing a function with 1_A yields the same function, except only for the detail that now we have to be careful about domains: $f \circ 1_A$ makes sense for functions $f : A \to B$ (for any B), whereas $1_A \circ f$ makes sense for functions $f : B \to A$ (for any B). The precise result is Theorem 99 below.

Theorem 99. *If A, B are sets, then*

$$f \circ 1_A = f \qquad \text{if} \qquad f : A \to B\,, \tag{22.513}$$

$$1_A \circ f = f \qquad \text{if} \qquad f : B \to A\,. \tag{22.514}$$

Proof. Suppose $f : A \to B$. Since $1_A : A \to A$, it follows that $f \circ 1_A : A \to B$. So both f and $f \circ 1_A$ have domain A.

To prove that the functions $f \circ 1_A$ and f are equal it suffices, according to Theorem 92, to prove that they have the same domain and that they have the same value for every x in that domain.

We already know that $f \circ 1_A$ and f have the same domain, because both have domain A. If $x \in A$, then

$$(f \circ 1_A)(x) = f(1_A(x)) = f(x)\,,$$

So $(f \circ 1_A)(x) = f(x)$. This completes the proof that $\boxed{f \circ 1_A = f}$.

Now let us suppose that $f : B \to A$. Since $1_A : A \to A$, it follows that $1_A \circ f : B \to A$. So both f and $1_A \circ f$ have domain B.

To prove that the functions $1_A \circ f$ and f are equal it suffices, according to Theorem 92, to prove that they have the same domain and that they have the same value for every x in that domain.

We already know that $1_A \circ f$ and f have the same domain, because both have domain B. If $x \in B$, then

$$(1_A \circ f)(x) = 1_A(f(x)) = f(x)\,,$$

So $(1_A \circ f)(x) = f(x)$, and this completes the proof that $\boxed{1_A \circ f = f}$.**Q.E.D.**

Theorem 100. *If A is a set, then 1_A is a bijection from A to A.*

Proof. **YOU DO IT.**

Problem 116. ***Prove*** Theorem 100. □

22.3.4 The inverse of a relation

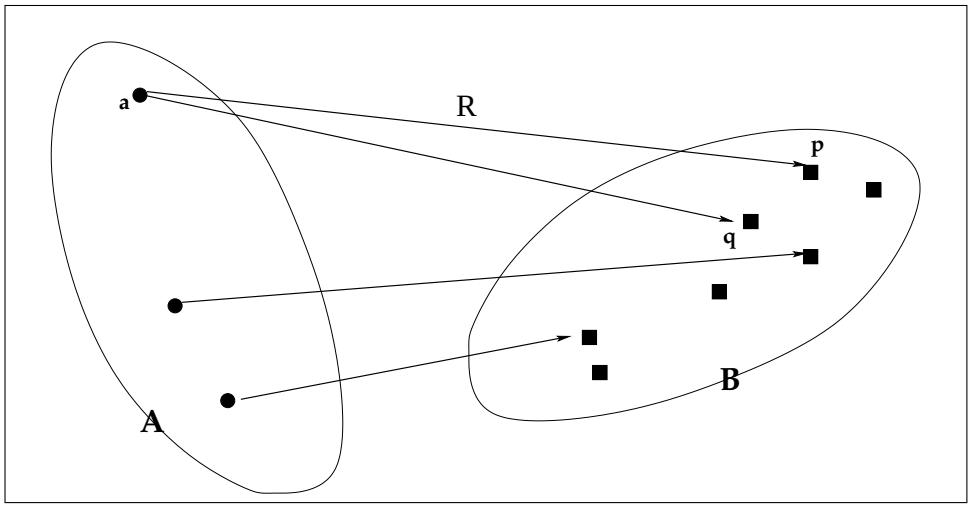
If you think of a relation R as a set of arrows (or of ordered pairs, which amounts to the same thing), then it is clear that can define another relation R^1 by just reversing the arrows of R: for every arrow of R going from a point x to a point y, we put in R^{-1} an arrow going from y to x. (Or, in terms of ordered pairs: R^{-1} consists of all the pairs (u, v) such that $(v, u) \in R$.)

Definition 70. *If R is a relation, then the <u>inverse</u> of R is the relation R^{-1} given by*

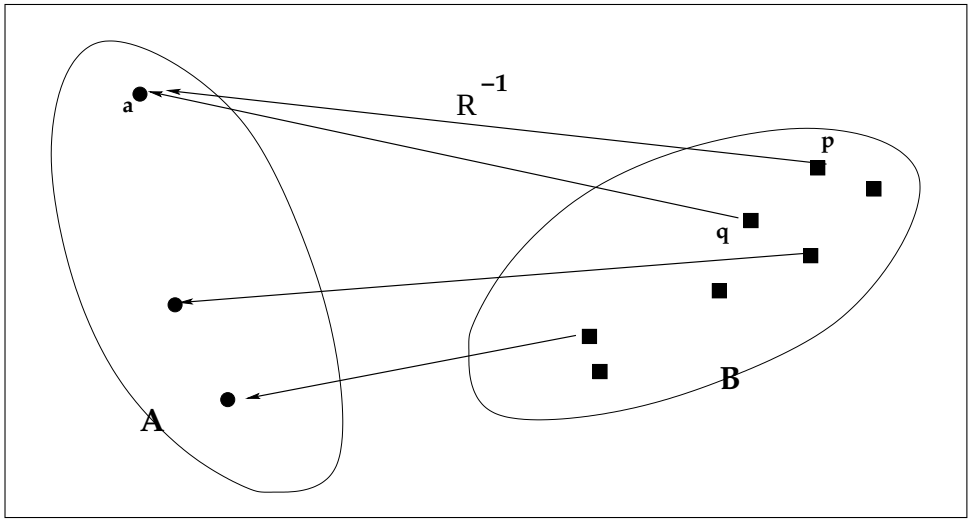
$$R^{-1} = \{(u, v) : (v.u) \in R\}\,. \tag{22.515}$$

Now that we know what the inverse of a relation is, it is clear that various properties of R correspond to properties of R^{-1}.

For example, consider the following relation (i.e., set of ordered pairs, or set of arrows) R:



and its inverse R^{-1}:



We see that R ***is not a function***, because there is one point a which is the source of two different arrowd ending at two different points$p$, q. (So R violates the unique output property: the input u gives rise to two different outputs, p and q.)

What does this tell us about R^{-1}? It tells us that R^{-1} is not one-to-one, because there are two different squares (i.e, inputs for R^{-1}) that are sources of arrows going to the same point. (Reason: both p and q are sent by R^{-1} to the same point a.)

This tells us that if R is a function then R^{-1} should be one-to-one, and if R is one-to-one then R^{-1} should be a function. In particular, we have:

Theorem 101. *If f is a function, then*

- *the relation f^{-1} is a function if and only if f is one-to-one,*
- *if f is one-to-one, then the domain of f^{-1} is the range of f.*

Proof. **YOU DO THIS.**

Problem 117. ***Prove*** Theorem 101. □

22.3.5 The inverse of the inverse

Theorem 102. *If R is a relation, then the inverse of the inverse of R is R. That is,*

$$R = \left(R^{-1}\right)^{-1}.$$

Proof. **YOU DO THIS.**

22.3.6 The inverse of a bijection

Theorem 103. *If A, B are sets, and f is a bijection from A to B, then f^{-1} is a bijection from B to A.*

Proof. We use Theorem 101 repeatedly.

Since f is a bijection, f is one-to-one, so by Theorem 101, f^{-1} is a function. Since the inverse of f^{-1} is f, and f is a function it follows from Theorem 101 that f^{-1} is one-to-one.

Since f is a function from A onto B, the range of f is B, so by Theorem 101 the domain of f^{-1} is B.

Since f^{-1} is a one-to-one, Theorem 101 tells that the domain of $(f^{-1})^{-1}$ is the range of f^1. But $(f^{-1})^{-1} = f$, so the domain of F is the range of f^{-1}. But the domain of f is A, so the range of f^{-1} is A, which menas that f^{-1} is a function onto A.

So $f^{-1} : B \to A$, f^{-1} is one-to-one, and f^{-1} is onto A. So f^{-1} is a bijection from B to A. **Q.E.D.**

22.3.7 Some problems

Problem 118. ***Prove*** Theorem 102.

HINT: This is trivial. The proof should be no more than one or two lines. □

Problem 119. ***Prove*** that if $f : A \longrightarrow B$, then f is a bijection from A to B if and only if the following is true:

(#) There exists a function $g : B \longrightarrow A$ such that $g \circ f = 1_A$ and $f \circ g = 1_B$. □

Problem 120. ***Prove or disprove*** each of the following statements. (NOTE; two of the statetents are true; the other four are false.)

1. If A, B, C are sets, $f : A \to B$ and $g : B \to C$ are functions, and $g \circ f$ is one-to-one, then f is one-to-one.
2. If A, B, C are sets, $f : A \to B$ and $g : B \to C$ are functions, and $g \circ f$ is one-to-one, then g is one-to-one.
3. If A, B, C are sets, $f : A \to B$ and $g : B \to C$ are functions, and $g \circ f$ is onto C then f is onto B.
4. If A, B, C are sets, $f : A \to B$ and $g : B \to C$ are functions, and $g \circ f$ is onto C, then g is onto to C.
5. If A, B, C are sets, $f : A \to B$ and $g : B \to C$ are functions, and $g \circ f$ is a bijection from A to C, then g is a bijection from B to C.
6. If A, B, C are sets, $f : A \to B$ and $g : B \to C$ are functions, and $g \circ f$ is a bijection from A to C, then f is a bijection from A to B.

23 Cardinality of sets

23.1 Sets with the same cardinality

If you look at the picture of a bijection on page 451, you can see right away, ***without having to count them***, that the number of squares is exactly the same as the number or circles. This is the crucial insight that leads to the following definition:

Definition 71. *Let A, B be sets. We say that*

- *B has the same number of members as A,*

or that

- *B has the same cardinality as A,*

and write

$$\mathrm{card}(A) = \mathrm{card}(B)\,, \tag{23.516}$$

if there exists a bijection from A to B,

Some authors call two sets "equivalent" if they have the same cardinality in the sense of our defintion. I do not like this because in mathematics there are already too many different meanings of thw word "equivalent" and I do not want to add one more meaning to the list.

Other authors use the word "equipotent", and you are welcome to use it if you wish. I just do not like it so I will not use it.

But it is true that "having the same number of members" is an equivalence relation, in the sense of the following theorem:

Theorem 104. *Let A, B, C be sets. Then:*

1. *A has the same cardinality as A;*
2. *if B has the same cardinality as A, then A has the same cardinality as B;*
3. *if B has the same cardinality as A, and C has the same cardinality as B, then C has the same cardinality as A,*

Proof. To prove that A has the same cardinality as A, we need a bijection from A to A. But we already know such a bijection, namely, the identity function 1_A. We proved in Theorem 100 that 1_A is a bijection from A to A, so $\boxed{A \text{ has the same cardinality as } A}$.

Now assume that B has the same cardinality as A. Then there exists a bijection f from A to B. And Theorem 103 tells us that f^{-1} is a bijection from B to A, so $\boxed{A \text{ has the same cardinality as } B}$.

Finally, assume that B has the same cardinality as A and C has the same cardinality as B. Then there exist a bijection f from A to B and a bijection g from B to C Theorem 98 then tells us that $g \circ f$ is a bijection from A to C. So $\boxed{C \text{ has the same cardinality as } A}$. **Q.E.D.**

23.2 Finite sets

23.2.1 An important notational convention: the sets $\mathbb{N}_k$

In what follows we will be making lots of statements about "the natural numbers $1, 2, \ldots, k$", that is "all the natural numbers j such that $j \leq k$". So it will be convenient to give a name to the set of all such js.

THE SETS $\mathbb{N}_k$ (A.K.A. $\{1, 2, \ldots, k\}$)

The expression "$\mathbb{N}_k$" stands for the set of all natural numbers that are less than or equal to k. That is,

$$\mathbb{N}_k = \{n \in \mathbb{N} : n \leq k\}. \tag{23.517}$$

Another notation often used for this set is "$\{1, \ldots, k\}$", or "$\{1, 2, \ldots, k\}$".
We will use "$\mathbb{N}_k$" when k is a natural number, and also when $k = 0$. (So $\mathbb{N}_k$ makes sense when $k \in \mathbb{N} \cup \{0\}$.)
Naturally, for $n = 0$ the set defined by (23.517) has no members, because there are no natural numbers k such that $k \leq 0$. So $\mathbb{N}_0 = \emptyset$.
Here are other examples:

$$\begin{array}{lll} \mathbb{N}_1 = \{1\}, & \mathbb{N}_2 = \{1, 2\}, & \mathbb{N}_3 = \{1, 2, 3\}, \\ \mathbb{N}_4 = \{1, 2, 3, 4\}, & \mathbb{N}_5 = \{1, 2, 3, 4, 5\}, & \mathbb{N}_6 = \{1, 2, 3, 4, 5, 6\}, \end{array}$$

Then

$$j \in \mathbb{N}_k$$
is just another way of saying "$j \in \mathbb{N}$ and $j \leq k$".

23.2.2 Finite lists

Definition 72.

- *A function whose domain is the set $\mathbb{N}_n$ for some nonnegative integer*[119] *is called a finite list of length n.*
- *If f is a finite list of length n and $k \in \mathbb{N}_n$, then $f(k)$ is the k-th entry of the list.*
- *If $f : \mathbb{N}_n \to A$ (so that every entry $f(k)$ of the finite list f is in A), then f is said to be a list of members of A.*
- *If $f : \mathbb{N}_n \to A$ and f is onto A (so that every entry $f(k)$ of the finite list f is in A and every member a of A occurs (in the list, in the sense that $a = f(k)$ for some $k \in \mathbb{N}_n$) then f is said to be a list of **all** the members of A.*
- *If $f : \mathbb{N}_n \to A$ and f is one-to-one (so that f has "no repeated entries", that is, $f(j)$ and $f(k)$ are never equal if $j \neq k$) then f is said to be a list of members of A without repetition.* □

Example 115. Let A be the set of all U.S. presidents from George Washington to Donald Trump. Since Donald Trump is the 45-th president, we can define a function $f : \mathbb{N}_{45} \to A$ by letting

$$f(k) = \text{the } k\text{−th U.S. president, for } l \in \mathbb{N}_{45}\,.$$

So, for example,

$$\begin{aligned} f(1) &= \text{George Washington}\,, \\ f(2) &= \text{John Adams}\,, \\ f(3) &= \text{Thomas Jefferson}\,, \\ &\cdots \\ f(16) &= \text{Abraham Lincoln}\,, \\ &\cdots \\ f(44) &= \text{Barack Obama}\,, \\ f(45) &= \text{Donald Trump}\,. \end{aligned}$$

Then f is a finite list of all the U.S. Presidents. (That is, $f : \mathbb{N}_{45} \to A$ and f is onto A.)

But f is not a one-to-one function (that is, f is not a list without repetitions), because Grover Cleveland is both the 22nd and the 24th U.S. president, so $f(22) = f(24)$. □

[119] "Nonnegative integer" means "natural number or zero". So the empty ser $\emptyset$ is a finite list of length 0, because 0 is a nonnegative integer, and $\mathbb{N}_0 = \emptyset$.

Theorem 105. *If A is a set, n is a nonnegative integer, and f is a list of length n of all the members of A (that is, $f : \mathbb{N}_n \to A$ and f is onto A), then there exist a nonnegative integer m such that $m \leq n$, and a list g of length m of all the members of A such that g is without repetition (that is, $g : \mathbb{N}_m \to A$, g is onto A, and g is one-to-one).*

Proof. **YOU DO THIS.**

Problem 121. ***Prove*** Theorem 105.

HINT: Eliminate the repetitions that occur in f, one at a time, until there are none left. Read Example 116 below to see how to eliminate one repetition. Then, in your proof, you have to elimeinate all the repetitions, one at a time. This means that you will have to do a proof by induction or by well-ordering. □

Example 116. In Example 115 we saw how to write a list f of length 45 of all the U.S. presidents. But this list is not without repetition (that is, f is not a one-to-one function) because $f(22) = f(24)$.

How can we create a list $g$of all the U.S. presidents without repetitions? The trick is to eliminate the repetition. Here is how:

Define $g : \mathbb{N}_{44} \to A$ by letting

$$g(k) = \begin{cases} f(k) & \text{if} \quad k \leq 23 \\ f(k+1) & \text{if} \quad k > 23 \end{cases} .$$

(In other words: from $k = 1$ up to $k = 23$ we don't change anything, so $g(k)$ is the same as $f(k)$. But for $k = 24$ we start changing things: we do not let $g(24)$ be $f(24)$, because of we did that we would get $g(24) = g(22)$, and there would be a repetition, What we do instead is ***skip over*** Grover Cleveland, and let $g(24)$ be $f(25)$ (that is we let $g(24)$ =William McKinley). And the we go on: $g(25)$ is $f(26)$, $g(26)$ is $f(27)$, and so on.)

Then g iss a list without repetitions of all the U.S. presidents. (That is, g is a one-to-one function from $\mathbb{N}_{44}$ onto A. So g is a bijection from $\mathbb{N}_n$ to A.) □

23.2.3 Finite sets and their cardinality

Definition 73. *If n is a nonnegative integer, then a set A is a <u>set with n members</u> (or <u>a set of cardinality n</u>) if there exists a list of length n of all the members of A without repetiotions.*

Using function langauge, we can say the same thing as follows:

if $n \in \mathbb{N} \cup \{0\}$, then A is a set with n members (or "A is a set of cardinality n") if there exists a bijection from $\mathbb{N}_n$ to A.

And the we can define the concepts of "finite set" and "infinite set".

Definition 74. *A set A is finite if it is a set of cardinality n for some nonnegative integer n.*
Equivalently, a set A is finite if for some nonnegative integer n here exists a bijection from $\mathbb{N}_n$ to A.

Definition 75. *A set A is infinite if it is not a finite set.*

Example 117.

1. The empty set is a set of cardinality 0. (Proof: as shown in Theorem 96, the empty set is a bijection from $\mathbb{N}_0$ to the empty set.)
2. Any singleton $\{a\}$ is a set of cardinality 1. (Proof: if $A = \{a\}$, define $f : \mathbb{N}_1 \to A$ by letting $f(1) = a$. Then f is a bijection from $\mathbb{N}_1$ to A.)
3. An undordered pair $\{a, b\}$ is a set of cardinality 2 if and only if $a \neq b$. (Proof: **YOU DO IT**.)
4. An undordered triple $\{a, b, c\}$ is a set of cardinality 3 if and only if $a \neq b$, $a \neq c$, and $b \neq c$. (Proof: **YOU DO IT**.)
5. If $n \in \mathbb{N} \cup \{0\}$, then the set $\mathbb{N}_n$ is a set of cardinality n. (Proof: the function $1_{\mathbb{N}_n}$ is a bijection from $\mathbb{N}_n$ to $\mathbb{N}_n$.)

Problem 122. Do the two proofs of parts 3 and 4 of Example 117. □

23.2.4 Can we talk about *the* cardinality of a finite set? The fundamental theorem of finite set cardinality theory

Now that we know what a finite set is, we would like to go further and give the following definition: *If A is a finite set and f is a bijection from $\mathbb{N}_n$ to A for some nonnegative integer n, then the number n is said to be the number of members of A, or the cardinality of A.*

But there is a problem. Suppose the following was possible, for some finite set A:

1. m and n are nonnegative integers,
2. there exists a bijection f from $\mathbb{N}_m$ to A,
3. there exists a bijection g from $\mathbb{N}_n$ onto A,
4. $m \neq n$

(For example, there could exist bijections from $\mathbb{N}_{10}$ to A and from $\mathbb{N}_{12}$ to A.)

If this happened, then we would not know which number should be called "***the*** cardinality of A". We would have to talk about "cardinalities of A", accepting that a finite set can have several different cardinalities, in the same way as, for example, an integer can have several factors, and several multiples, so we do not say "6 is *the* factor of 30", or "30 is *the* multiple of 6"; we say "6 is *a* factor of 30" and "30 is *a* multiple of 6".

Fortunately, this problem does not occur. The cardinality of a finite set is unique, so we *can* talk about "the" cardinanlty of a finite set. This is so because of the following theorem:

Theorem 106. *Assume that A is a set and m and n are nonnegative integers such that there exists a bijection f from $\mathbb{N}_m$ to A, and there exists a bijection g from $\mathbb{N}_n$ to A. Then $m = n$.*

Thanks to Theorem 106, if I know that a set A if of cardinality n, then I can say that n is *the* cardinality (or *the* number of members) of A.

Proof of Theorem 106. The key point of the proof is the exchange lemma that we proved earlier as Theorem 97. We will ue the exchange lemma to prove Lemma 4, which is essentially the inductive step in the proof by induction of Lemma 3. which easily implies our result.

Let us assume that f is a bijection from $\mathbb{N}_m$ to A, g is a bijection from $\mathbb{N}_n$ to A. We want to prove that $m = n$.

Since g is a bijection from $\mathbb{N}_n$ to A, Theorem 103 tells us that g^{-1} is a bijection from A to $\mathbb{N}_n$.

Now we have bijections

$$f : \mathbb{N}_m \to A\,, \quad g^{-1} : A \to \mathbb{N}_n\,.$$

And then Theorem 98 tells us that the composite function $h : \mathbb{N}_m \to \mathbb{N}_n$, defined by $h = g^{-1} \circ f$, is a bijection from $\mathbb{N}_m \to \mathbb{N}_n$.

So all we need to do prove the following lemma:

Lemma 3. *If $m \in \mathbb{N} \cup \{0\}$, $n \in \mathbb{N} \cup \{0\}$, and there exists a bijection from $\mathbb{N}_m$ to $\mathbb{N}_n$, then $m = n$.*

To prove Lemma 3, we use the exchange lemma, i.e., Theorem 97.

Suppose there exists a bijection f from $\mathbb{N}_m$ to $\mathbb{N}_n$. Then by Theorem 97, *there exists a bijection k from $\mathbb{N}_m$ to $\mathbb{N}_n$ that has the additional property that $f(m) = n$.*

But then, if we remove the pair (m, n) from k, we get a bijection from $\mathbb{N}_{m-1}$ to $\mathbb{N}_{n-1}$. So we have proved[120]

Lemma 4. *If $m \in \mathbb{N}$, $n \in \mathbb{N}$, and there exists a bijection from $\mathbb{N}_m$ to $\mathbb{N}_n$, then there exists a bijection from $\mathbb{N}_{m-1}$ to $\mathbb{N}_{n-1}$.*

Once we have Lemma 4, we can do a proof of Lemma 3 by induction or by well-ordering.

I will give you the proof using well-ordering.

Call a nonnegative integer n "bad" if there exist a nonnegative integer m such that $m \neq n$ and there exists a bijection from $\mathbb{N}_m$ to $\mathbb{N}_n$.

We want to prove that there are no bad nonnegative integers. In other words, if we let B be the set of all bad nonngative integers, we want to prove that B is empty.

We first prove that 0 is not bad.

Proof that 0 is not bad.

- Assume that $\boxed{\text{0 is bad}}$.
- Since 0 is bad, there exist a nonnegative integer m such that $m \neq 0$, and a bijection f from $\mathbb{N}_m$ to $\mathbb{N}_0$.

[120] Why have I suddenly switched from "$m \in \mathbb{N} \cup \{0\}$ and $n \in \mathbb{N} \cup \{0\}$" to "$m \in \mathbb{N}$ and $n \in \mathbb{N}$"? That's because if $m = 0$ or $n = 0$ then k is empty so the pair (m, n) is not in k and cannot be removed from k.

- Since $m \neq 0$, $m \in \mathbb{Z}$, and $m \geq 0$, it follows that $m \in \mathbb{N}$, sp $1 \in \mathbb{N}_m$.
- On the other hand, since f is bijection from $\mathbb{N}_m$ to $\mathbb{N}_0$, the domain of f is the set $\mathbb{N}_m$, so $1 \in \text{Dom}(f)$.
- Then $f(1) \in \mathbb{N}_0$, so $\boxed{\mathbb{N}_0 \neq \emptyset}$.
- But $\boxed{\mathbb{N}_0 = \emptyset}$.

So the assumption that 0 is bad has led us to a contradiction. Hence $\boxed{0 \text{ is not bad}}$.

We now prove that B is empty, and do it by contradiction.

- Assume that $\boxed{B \neq \emptyset}$.
- Then B is a nonempty subset of $\mathbb{Z}$, and B is bounded below because every member of B is ≥ 0.
- So by the well-ordering principle, B has a smallest member b.
- Then $b \in \mathbb{Z}$, $b \geq 0$, and b is bad.
- So in particular $b \neq 0$, because we already know that 0 is not bad.
- Then $b \in \mathbb{N}$.
- Since b is bad, there exist a nonnegative integer m such that $m \neq b$, and a bijection f from $\mathbb{N}_m$ to $\mathbb{N}_b$,
- Since $b \in \mathbb{N}$, $b \geq 1$, so $1 \in \mathbb{N}_b$.
- Since f is onto $\mathbb{N}_b$, and 1 belongs to $\mathbb{N}_b$, we may pick $x \in \text{Dom}(f)$ such that $f(x) = 1$.
- Then $\text{Dom}(f) \neq \emptyset$, so $\mathbb{N}_m \neq \emptyset$, and then $m \neq 0$.
- Since $m \in \mathbb{Z}$, $m \geq 0$, and $m \neq 0$, it follows that $m \in \mathbb{N}$.
- Since f is a bijection from $\mathbb{N}_m$ to $\mathbb{N}_b$, and both m and b are natural numbers, we can applu Lemma 4 and conclude that there exists a bijection g from $\mathbb{N}_{m-1}$ to $\mathbb{N}_{b-1}$.
- But $m \neq b$, so $m - 1 \neq b - 1$.
- So $\boxed{b - 1 \text{ is bad}}$.

- But $\boxed{b-1 \text{ is not bad}}$, because we are assuming that b is the smallest bad integer.

So the assumption that B is nonempty has led us to a contradiction.

Hence B is empty, and our proof of Lemma 3 is complete.

End of the proof of Theorem 106. As explained before, if f is a bijection from $\mathbb{N}_m$ to A and g is a bijection from $\mathbb{N}_n$ to A, then $g^{-1} \circ f$ is a bijection from $\mathbb{N}_m$ to $\mathbb{N}_n$. Then Lemma 3 telts that $m = n$, and proof of Theorem 106 is complete. **Q.E.D.**

Problem 123. We have given a proof of Lemma 3 using well-ordering. ***Prove*** Lemma 3 using induction.

You are allowed to use Lemma 4. □

23.2.5 Definition of "cardinality" of a finite set

Now that we have proved Theorem 106, we can talk about *the* cardinality (or *the* number of members) of a finite set A.

Definition 76. *Let A be a finite set. Then the nonnegative integer n such that*

(*) there exists a bijection from $\mathbb{N}_n$ to A,

is called the cardinality of A, or the number of members of A. (The number n exists because A is finite, and is unique thanks to Theorem 106.)

We write "$\text{card}(A)$*" to denote the cardinality of A.* □

Problem 124. ***Prove*** that:

1. If A is a finite set, B is a set, and there exisrs a bijection from B to A, then B is finite and $\text{card}(B) = \text{card}(A)$.

2. If A and B are finite sets, and $\text{card}(A) = \text{card}(B)$, then there exisrs a bijection from B to A. □

23.2.6 A trivial but important lemma

The following lemma is very obvious but, as all obvious things is this game, needs proof. The lemma says that if you have a set with n members and remove one member then you are left with a set with $n-1$ members.

Lemma 5. *If A is a finite set, and a is a member of A, then the set $A - \{a\}$ is finite and has cardinality* $\text{card}(A) - 1$.

Proof. Let $n = \text{card}(A)$. Then we may pick a bijection f from $\mathbb{N}_n$ to A. Thanks to the exchange theorem (i.e., Theorem 97) there exists a bijection g from $\mathbb{N}_n$ to A such that $g(n) = a$.

Let $h = g - \big\{(n, a)\big\}$. (That is, h is the set of ordered pairs obtained from g by removing the pair (n, a).)

Then h is a bijection from $\mathbb{N}_{n-1}$ to $A - \{a\}$. (This is easy to prove. **YOU SHOULD PROVE IT.**)

So $A - \{a\}$ is a finite set and $\text{card}(A - \{a\}) = \text{card}(A) - 1$. **Q.E.D.**

23.2.7 Subsets of a finite set

Definition 77. *A proper subset of a set A is a subset B of A such that $B \neq A$.* □

Theorem 107. *Let A be a finite set, and let B be a proper subset of A. Then B is finite and* $\text{card}(B) < \text{card}(A)$.

Proof. We will use induction.

Let $P(n)$ be the predicate

$P(n)$ *If A is a finite set with cardinality n, then every proper subset B of A is finite and has cardinality $< n$.*

We prove that $P(n)$ is true for every nonnegative integer n, by induction on n starting at $n = 0$.

Basis step. $P(0)$ is true because, if A if a finite set with cardinality 0, then A must be the empty set, so A has no proper subsets, so the statement "if B is a proper subset of A then B is finite and cars$(B) < 0$" is vacuously true.

Inductive step. Assume $P(n)$ is true.

We want to prove $P(n + 1)$. That is, we want to prove that

(#) If A is a finite set with cardinality $n + 1$, then every proper subset of A is finite and has cardinality $< n$.

Let A be a finite set with cardinality $n + 1$, and let B be a proper subset of A. We want to prove that

(##) B is finite and $\text{card}(B) < n + 1$.

Since B is a proper subset of A, $B \neq A$, so there exists $a \in A$ such that $a \notin B$. (Reason: if every member of A was in B, it would follow that $A \subseteq B$. Since $B \subseteq A$, this would imply $B = A$, contradicting our assuption that $B \neq A$.) Let $A' = A - \{a\}$. Then $B \subseteq A'$.

By Lemma 5, A' is a finite set with cardinality n.

The set B is either equal to A' or a proper subset of A'.

If $B = A'$, then B is finite and $\text{card}(B) = n$, so (##) holds.

If B is a proper subset of A' then, by the inductive hypothesis $P(n)$, B is finite and $\text{card}(B) < n$, so *a fortiori* (##) holds.

So (##) holds in both cases. Therefore $P(n+1)$ is true.

This completes the induction, and then the proof of Theorem 107.**Q**.**E**.**D**.

The following is a slightly stronger version of Theorem 107, in which the subset B is not required to be proper.

Theorem 108. *Let A be a finite set, and let B be a subset of A. Then*

1. *B is finite,*
2. $\text{card}(B) \leq \text{card}(A)$*,*
3. *if B is a proper subset of A then* $\text{card}(B) < \text{card}(A)$.

Proof. Let A be a finite set with cardinailty n, and let B be a subset of A. If B is proper, them Theorem 107 tells us that B is finite and $\text{card}(B) < n$.

If B is not proper, then $B = A$, so B is finite and $\text{card}(B) = n$.

So our desired conclusion holds in both cases. **Q**.**E**.**D**.

23.2.8 The Dirichlet pigeonhole principle

We are going to use another very important result.

Theorem 109. *Assume that A, B are finite sets and $f : A \to B$ is a one-to-one function. Then*

1. $\text{card}(A) \leq \text{card}(B)$*,*
2. $\text{card}(A) < \text{card}(B)$ *if and only if f is not onto B.*

Theorem 109 is known as the Dirichlet pigeonhole principle (DPHP), for the following reason.

Think of A as a set of pigeons, and B as a set of holes. The function f assigns a hole $f(p)$ to each pigeon p. The fact that f is one-to-one says that different pigeons go to different holes, i.e., that it does no happen that two different pigeons are assigned to the same hole. The theorem then says that

1. there are at least as many holes as there are pigeons,
2. the number of pigeons is equal to the number of holes if and only if every hole is occupied by a pigeon.

Here is another example: suppose that in a classroom there are m students and n seats, and each student is seated, and no seat has more than one student in it. Then the DPHP says that there are at least as many seats as there are students.

Proof of Theorem 109. **YOU DO IT.**

Problem 125. ***Prove*** Theorem 109.

HINT: Let C be the range of f, so C is a subset of B. Prove first that f is a bijection from A to C. Then use Theorems 107 and 108. □

Here is another version of the pigeonhole principle. This time, we look at the case when every hole is occupied by at least one pigeon. The conclusion in this case is that

1. there are at least as many pigeons as there are holes.
2. the number of pigeons is equal to the number of holes if and only if no hole is occupied by more than one pigeon.

Theorem 110. *Assume that A, B are finite sets and $f : A \to B$ is a function onto B. Then*

1. $\text{card}(A) \geq \text{card}(B)$,

2. $\text{card}(A) > \text{card}(B)$ *if and only if f is not one-to-one.*

Proof. **YOU DO IT.**

Problem 126. ***Prove*** Theorem 110.

HINT: Construct a one-to-one "hole-to-pigeon" function by picking for each hole h a pigeon p that occupies hole h. Show that this function is one-to-one and the use Theorem 109. □

Theorem 111. *Let A, B be finite sets. Then*

1. $\mathrm{card}(A) \leq \mathrm{card}(B)$ *if and only if there exists a one-to-one function from A to B,*
2. $\mathrm{card}(A) < \mathrm{card}(B)$ *if and only if there exists a one-to-one function from A to B but there does not exist a one-to-one function from B to B.*

Proof. **YOU DO IT.**

Theorem 112. *Let A, B be finite sets. Then* $\mathrm{card}(A) \leq \mathrm{card}(B)$ *if and only if there exists a function from B onto A.*

1. $\mathrm{card}(A) \leq \mathrm{card}(B)$ *if and only if there exists a function from B onto A,*
2. $\mathrm{card}(A) < \mathrm{card}(B)$ *if and only if there exists a function from B onto A but there does not exist a function from A onto B.*

Proof. **YOU DO IT.**

Problem 127. ***Prove*** Theorems 111 and 112. □

23.2.9 Unions of finite sets

Theorem 113. *Let A, B be disjoint*[121] *finite sets. Then $A \cup B$ is finite, and*

$$\mathrm{card}(A \cup B) = \mathrm{card}(A) + \mathrm{card}(B)\,. \tag{23.518}$$

Proof. **YOU DO IT.**

Problem 128. ***Prove*** Theorem 113.

HINT: You an do this by induction with respect to card(A)or card(B). Or you can do it directly, by combining a bijection f from $\mathbb{N}_m$ to A and a bijection g from $\mathbb{N}_n$ to B to construct a bijection h from $\mathbb{N}_{m+n}$ to $A \cup B$.□

Theorem 114. *Let A, B be finite sets. Then $A \cup B$ is finite, and*

$$\mathrm{card}(A \cup B) = \mathrm{card}(A) + \mathrm{card}(B) - \mathrm{card}(A \cap B)\,. \tag{23.519}$$

Proof. **YOU DO IT.**

[121] Two sets A, B are disjoit if $A \cap B = \emptyset$.

Problem 129. ***Prove*** Theorem 114.

HINT: Divide $A \cup B$ intro three sets C, D, E as follows: $C = \{x : x \in A \wedge x \notin B\}$, $D = \{x : x \in B \wedge x \notin A\}$, and $E = A \cap B$. Tnen C, D, F are finite by Theorem 108.

Also, $C \cap E = \emptyset$, $C \cup E = A$, $D \cap E = \emptyset$, $D \cup E = B$, $A \cap D = \emptyset$, $A \cup D = A \cup B$. □

23.2.10 Sets of subsets

Theorem 115. *Let A be a finite set. Then the power set $\mathcal{P}(A)$ is finite, and*

$$\operatorname{card}\Big(\mathcal{P}(A)\Big) = 2^{\operatorname{card}(A)}\,. \tag{23.520}$$

Proof. **YOU DO IT.**

Problem 130. ***Prove*** Theorem 115.

HINT: Do it by induction on $n = \operatorname{card}(B)$.

Fix $a \in A$, and let $A' = A - \{a\}$. Then the subsets of A are of two kinds: those that do not contain a as a member and those that do. The subsets of the first kind are exactly the subsets of A', so by the inductive hypthesis there are 2^{n-1} such sets. If $\mathcal{B}$ is the set of all the subsets of the second kind, then you should construct a bijection from $\mathcal{B}$ to $\mathcal{P}(A')$. Then $\mathcal{P}(A) = \mathcal{P}(A') \cup \mathcal{B}$, $\mathcal{P}(A') \cup \mathcal{B} = \emptyset$, and $\operatorname{card}(\mathcal{B}) = \operatorname{card}\big(\mathcal{P}(A')\big)$. □

23.2.11 Cartesian products of finite sets

Theorem 116. *Let A, B be finite sets. Then $A \times B$ is finite, and*

$$\operatorname{card}(A \times B) = \operatorname{card}(A) \cdot \operatorname{card}(B)\,. \tag{23.521}$$

Proof. **YOU DO IT.**

Problem 131. ***Prove*** Theorem 116.

HINT: Do it by induction on $m = \operatorname{card}(B)$.

Prove that if $b \in B$, and $B' = B - \{b\}$, then

$$A \times B = (A \times B') \cup (A \times \{b\})\,, \text{ and } (A \times B') \cap (A \times \{b\}) = \emptyset\,.$$

and use this in your inductive argument. □

23.3 Infinite sets

We recall the definition of "infinite set": *A set is infinite if it is not a finite set.*

Theorem 117. *The set* $\mathbb{N}$ *of all natural numbers is infinite.*

Idea of the proof. If $\mathbb{N}$ was finite, then it would have some cardinality $n \in \mathbb{N} \cup \{-\}$, and this means that we can put a pigeon for each $k \in \mathbb{N}$ and fit all these pigeons in n holes. But $\mathbb{N}$ has at least m members, for every m. So for every m we can fit m pigeons in n holes. So $m \geq n$ by the Dirichlet pigeonhole principle. So the number n is greater than every natural number, But such an n cannot exist.

Now let us write this down in mathematical language.

Proof. Suppose $\mathbb{N}$ was finite.

Then there exist a natural number n and a bijection f from $\mathbb{N}_n$ to $\mathbb{N}$.

The inverse function $g = f^{-1}$ is then a bijection from $\mathbb{N}$ to $\mathbb{N}_n$.

Now let m be an arbitrary natural number.

Define a function g_m, with domain $\mathbb{N}_m$, by letting

$$g_m(k) = g(x) \text{ for } k \in \mathbb{N}_m\,.$$

Then $g_m : \mathbb{N}_m \to \mathbb{N}$ and g_m is one-to-one. (Proof: if $k_1, k_2 \in \mathbb{N}_m$ and $g_m(k_1) = g_m(k_2)$ then $g(k_1) = g(k_2)$, so $k_1 = k_2$ because g is one-to-one.

Furthermore, $g : \mathbb{N}_m \to \mathbb{N}_n$. It the follows from the Dirichlet pigeonhole principle (Theorem 109) that $n \geq m$.

Since n was an arbitrary real number, we have proved that

$$(\forall m \in \mathbb{N}) n \geq m\,. \tag{23.522}$$

So we have found a natural number n which is larger than every natural number.

But we know that such a number cannot exist. (That is, (23.522) is impossible. This is easily seen as follows: if (23.522) was true for some $n \in \mathbb{N}$, then we could specialize to $m = n + 1$ and conclude that $n \geq n + 1$. But $n < n + 1$, so we have arrived at a contradiciton.)

So the assumption that $\mathbb{N}$ is finite has led us to a contradiction.**Q.E.D**.

Theorem 118. *If A, B are sets, f is a bijection from A to B, and A is infinite, then B is infinite.*

Proof. **YOU DO THIS.**

Problem 132. ***Prove*** theorem 118. □

Theorem 119. *If A is a set, B is a subset of A, and B is infinite, then A is infinite.* □

Proof. **YOU DO THIS.**

Problem 133. ***Prove*** theorem 119. □

Theorem 120. *Define a function f from $\mathbb{N}$ to $\mathbb{Z}$ as follows:*

$$f(n) = \begin{cases} \frac{n}{2} & \text{if} \quad n \in \mathbb{N} \text{ and } n \text{ is even} \\ \frac{1-n}{2} & \text{if} \quad n \in \mathbb{N} \text{ and } n \text{ is odd} \end{cases} . \tag{23.523}$$

Then f is a bijection from $\mathbb{N}$ to $\mathbb{Z}$.

Proof. **YOU DO THIS.**

Problem 134. ***Prove*** Theorem 120. □

Problem 135. Let $\mathcal{E}_{>0}$ be the set of all even natural numbers. That is,

$$\mathcal{E}_{>0} = \{n \in \mathbb{N} : 2|n\}, \tag{23.524}$$

and let $\mathcal{E}$ be the set of all even integers, so

$$\mathcal{E} = \{n \in \mathbb{Z} : 2|n\}, \tag{23.525}$$

Construct bijections from $\mathbb{N}$ to $\mathcal{E}_{>0}$ and from $\mathbb{N}$ to $\mathcal{E}$. □

NOTE: If f is a function whose domain in $\mathbb{N} \times \mathbb{N}$, the set of all order pairs (m, n) of natural numbers, and u is a member of the domain of f, we have to write $f(u)$ for the value of f at u. But u is itself an ordered pair (m, n), so we should write $f\big((m, n)\big)$ for rhe value of f at u. i.e., at (m, n). But we are going to follow the starnard practice of omitting one pair of parentheses and just write "$f(m, n)$", instead of "$f\big((m, n)\big)$".

Theorem 121. *Define a function f from $\mathbb{N} \times \mathbb{N}$ to $\mathbb{N}$ as follows:*

$$f(m, n) = 2^{m-1}(2n - 1) \text{ for} \tag{23.526}$$

Then f is a bijection from from $\mathbb{N} \times \mathbb{N}$ to $\mathbb{N}$.

Proof. **YOU DO THIS.**

Problem 136. ***Prove*** Theorem 121. □

Problem 137. ***Construct*** a partition of $\mathbb{N}$ into infinitely many infinite sets. (The definition of "partition" is given in a previous homework.)

HINT: Using the result of Theorem 121, this should juat require two or three lines. □

Theorem 122. *There exists a bijection from $\mathbb{N}$, the set of all naatural numbers, to $\mathbb{Q}$, the set of all rational numbers.*

Proof. This proof is done in the book. You should look it up there.

Problem 138. ***Prove*** that there exists a bijection from $\mathbb{N}$ to $\mathbb{Q} \times \mathbb{Q}$. □

23.3.1 Countable sets

Definition 78.

- *A set A is <u>countable</u> if there exists a one-to-one function $f : A \to \mathbb{N}$.*
- *A set A is <u>counatbly infinite</u> if it is countable and infinite.* □

It is clear that

1. $\mathbb{N}$ *is countable.* (Proof: the function $1_{\mathbf{N}}$ is a bijection from $\mathbb{N}$ to $\mathbb{N}$, so in particular it is a one-to-one function from $\mathbb{N}$ to $\mathbb{N}$.
2. *Every subset of a countable set is countable.* (Proof: Let A be countable, and let B be a subset of A. Let $f : A \to \mathbb{N}$ be a one-to-one function. Define $g : B \to \mathbb{N}$ by letting $g(x) = f(x)$ for each $x \in B$. Then g is one-to-one.)
3. *Every finite set is countable.*
4. $\mathbb{N}$ *is countably infinite.* (Proof: $\mathbb{N}$ is countable, and $\mathbb{N}$ is infinite because of Theorem 117.)
5. If A, B are sets and f is a bijection from A to B, then A is countably infinite if and only if B is.
6. $\mathbb{Z}$, $\mathbb{N} \times \mathbb{N}$, *and* $\mathbb{Q}$ *are countably infinite.* (Reason: we have already constructed bijections from $\mathbb{N}$ to $\mathbb{Z}$ and from $\mathbb{N}$ to $\mathbb{N}\times\mathbb{N}$, and Theorem 122 tells us that there exists a bijection from $\mathbb{N}$ to $\mathbb{Q}$.)

23.3.2 Do all infinite sets have the same cardinality?

The results of Theorems 120, 121 and problems 135 show that sets as diverse as $\mathbb{Z}$, $\mathcal{E}_{>0}$, $\mathcal{E}$, and $\mathbb{N} \times \mathbb{N}$, some of which (for example, $\mathbb{Z}$ and $\mathbb{N} \times \mathbb{N}$) appear to be much larger than $\mathbb{N}$, have in fact the same cardinality as $\mathbb{N}$.

Could it be that all infinite sets have the same cadinality as $\mathbb{N}$?

The asnwer is a resounding NO!!!. Here is a result which is very easy to prove but has momentous consequences.

Theorem 123 *(Cantor) If X is a set, then there does not exist a function from X onto the power set $\mathcal{P}(X)$.*

Proof. Suppose $f : X \to \mathcal{P}(X)$ is onto $\mathcal{P}(X)$.

Let

$$S = \{x \in X : x \notin f(x)\}.$$

Then S is a subset of X, so $S \in \mathcal{P}(X)$. Since f is onto $\mathcal{P}(X)$, we can pick $s \in X$ such that $f(s) = S$.

Let us prove that $s \in S$. Suppose $s \notin S$. Then $s \notin f(s)$, so s satisfies the membership criterion for S ("$x \notin f(x)$"). Therefore $s \in S$. But we are assuming that $s \notin S$, so we have reached a contradiciton. So the assumption that $s \notin S$ had led us to a contradiction, and then it follows that $s \in S$.

Now let us prove that $s \notin S$. Suppose $s \in S$. Then $s \in f(s)$, so s does not satisfy the membership criterion for S ("$x \notin f(x)$"). Therefore $s \notin S$. But we are assuming that $s \in S$, so we have reached a contradiciton. So the assumption that $s \in S$ had led us to a contradiction, and then it follows that $s \notin S$.

We have proved that $\boxed{s \in S \wedge s \notin S}$. So we have arrived at a contradiction. The contradiction resulted from assuing that there exists a function f from X onto $\mathcal{P}(X)$. Hence $\boxed{\text{a function } f \text{ from } X \text{ onto } \mathcal{P}(X) \text{ does not exists}}$. **Q.E.D.**

23.3.3 Consequences of Cantor's Theorem

Theorem 123 says that if X is any set then there does not exist a function from X onto $\mathcal{P}(X)$. So in particular there cannot exist a bijection from X to $\mathcal{P}(X)$.

This means that the sets X and $\mathcal{P}(X)$ do not have the same cardinality. Can we say that one of them is "larger" than the other one? The answer is "yes", $\mathcal{P}(X)$ definitely has a larger cardinality than X. But before we say that, we have to know what it means: ***what does it mean for a set A to have a larger cardinality than that of a set B?***

In order to asnwer that question, we first look at what happens for finite sets:

(I) First of all, if A and B are finite sets, then the cardinalities $\text{card}(A)$ and $\text{card}(B)$ are nonnegative integers, and we know what it means for a nonnegative integer to be larger than another nonnegative integer.

(II) Second, Theorem 30 tells us that, if A and B are finite sets, then

1. $\text{card}(A) \leq \text{card}(B)$ if and only if there exists a one-to-one function from A to B,
2. $\text{card}(A) < \text{card}(B)$ if and only if there exists a one-to-one function from A to B but there does not exist a one-to-one function from B to B.

The conditions of (II) make perfect sense for infinite sets as well, even though we do not know what "$\text{card}(A)$" means. (We have only defined what it means for two sets A, B to "have the same cardinality". This does not say that there is some object called "the cardinality of a set", and that two sets have the same cardinality if and only if that object is the same for both. For finite sets, we were able to define such an object, and it turned out to be a nonnegative integer. For infinite sets, this can be done too, but we have not done it yet, so at this point we do not know what "$\text{card}(A)$" means. All we know is what "$\text{card}(A) = \text{card}(B)$" means. And in the next section we are going to assign a menaning to "$\text{card}(A) \leq \text{card}(B)$", "$\text{card}(A) < \text{card}(B)$", "$\text{card}(A) \geq \text{card}(B)$", and "$\text{card}(A) > \text{card}(B)$", but we will still not know what "$\text{card}(A)$" means.

23.3.4 Comparing sizes of sets. The Cantor-Schroeder-Bernstein Theorem

We follow the idea of Theorem 30. (We could also follow the second one, and it works, but there are some complications.)

Definition 79. *Suppose that Let A, B are sets.*

- *We say that A has cardinality smaller than or equal to that of B, or that B has cardinality larger than or equal to that of A, and write*

$$\text{card}(A) \leq \text{card}(B)\,,$$

or

$$\text{card}(B) \geq \text{card}(A)\,,$$

if there exists a one-to-one function from A to B.

- *We say that A has cardinality strictly smaller than that of B, or that B has cardinality strictly larger than that of A, or that and write*

$$\text{card}(A) < \text{card}(B)$$

or

$$\text{card}(B) > \text{card}(A)\,,$$

if there exists a one-to-one function from A to B, but there does not exist a bijection from A to B. □

As an important example of the use of these definitions, we prove the following theorem.

Theorem 124 *(Cantor) If X is a set, then the cardinality of X is strictly smaller than the cardinality of the power set $\mathcal{P}(X)$. That is:*

$$\operatorname{card}(X) < \operatorname{card}\big(\mathcal{P}(X)\big)\,. \tag{23.527}$$

Proof. First, we show that $\operatorname{card}(X) \leq \operatorname{card}\big(\mathcal{P}(X)\big)$, by constructing a one-to-one function $f : X \to \mathcal{P}(X)$.

This is very easy: define $f : X \to \mathcal{P}(X)$ by letting

$$f(x) = \{x\} \qquad \text{for} \quad x \in X\,.$$

Then f is clearly one-to-one. (Proof: suppose $x_1, x_2 \in X$ and $f(x_1) = f(x_2)$; then $\{x_1\} = \{x_2\}$; since $x_2 \in \{x_2\}$, it follows that $x_2 \in \{x_1\}$; but x_1 is the only member of $\{x_1\}$; so $x_2 = x_1$.)

So we have constructed a one-to-one function from X to $\mathcal{P}(X)$.

On the other hand, Theorem 123 tells us that there does not exist a function from X onto $\mathcal{P}(X)$. In particular, there does not exist a bijection from X to $\mathcal{P}(X)$.

This proves that (23.527) holds. **Q.E.D.**

The relations "$\leq$", "$\geq$", "$<$", "$>$", should behave like their homonyms[122] the relations "$\leq$", "$\geq$", "$<$", "$>$", between real numbers.

[122] ***Homonyms*** are words or symbols that are spelled or written identically but have different meanings. For example, "crane", a mechanical lifting machine, and "crane", a bird, are homonyms. Mathematics is full of homonyms. For example, "$<$" as a binary relation between real numbers, and "$<$" as a binary relation between cardinalities, are different meanings of the symbol "$<$".

So, for example, we would expect that there is a true theorem about inequalities among cardinals corresponding to each of the following properties of inequalities between real numbers:

(R1) If $a, b \in \mathbb{R}$, then $a \leq b$ if and only if $b \geq a$.
(R2) If $a, b \in \mathbb{R}$, then $a < b$ if and only if $b > a$.
(R3) If $a, b, c \in \mathbb{R}$, $a \leq b$, and $b \leq c$, then $a \leq c$.
(R4) If $a, b \in \mathbb{R}$, then $a \leq b$ if and only if $a < b$ or $a = b$.
(R5) If $a, b \in \mathbb{R}$ and $a < b$, then $a \neq b$.
(R6) If $a, b \in \mathbb{R}$, $a \leq b$, and $b \leq a$, then $a = b$.
(R7) If $a, b, c \in \mathbb{R}$, $a \leq b$, $b \leq c$, and either $a < b$ or $b < c$, then $a < c$.
(R8) If $a, b \in \mathbb{R}$, then either $a \leq b$ or $b \leq a$.

It turns out that ***the analogues of the nine properties listed above are all true; some are trivially true, but others are not at all obvious and their proofs require a lot of work.***

Let us start with the obvious ones: the analogues of (R1), (R2), (R3), (R4), (R5) for cardinalities are trivially true. We say that in the following theorem:

Theorem 125. *Let A, B, C be sets. Then*

(C1) $\text{card}(A) \leq \text{card}(B)$ *if and only if* $\text{card}(A) \geq \text{card}(B)$.

(C2) $\text{card}(A) < \text{card}(B)$ *if and only if* $\text{card}(B) > \text{card}(A)$.

(C3) If $\text{card}(A) \leq \text{card}(B)$ *and* $\text{card}(B) \leq \text{card}(C)$ *then* $\text{card}(A) \leq \text{card}(C)$

(C4) $\text{card}(A) \leq \text{card}(B)$ *if and only if either* $\text{card}(A) < \text{card}(B)$ *or* $\text{card}(A) = \text{card}(B)$.

(C5) If $\text{card}(A) < \text{card}(B)$ *then it's not true that* $\text{card}(A) = \text{card}(B)$.

Proof. Statement (C1) is true because Definition 79 tells us "$\text{card}(A) \leq \text{card}(B)$" and "$\text{card}(B) \geq \text{card}(A)$" are two ways of writing the same thing.

Similarly, statement (C2) is true because Definition 79 tells us "$\text{card}(A) < \text{card}(B)$" and "$\text{card}(B) > \text{card}(A)$" are two ways of writing the same thing.

To prove (C3), assume that $\text{card}(A) \leq \text{card}(B)$ and $\text{card}(B) \leq \text{card}(C)$. This means, according to Definition 79, that there exist one-to-one functions $f : A \to B$ and $g : B \to C$. Then the composite $g \circ f$ is a function from A to C, and Theorem 94 tells us that $g \circ f$ is one-to-one. Hence $g \circ f$ is a one-to-one function from A to C. So $\text{card}(A) \leq \text{card}(C)$.

Next, we prove (C4). Wew have to prove that

$$\mathrm{card}(A) \leq \mathrm{card}(B) \Longleftrightarrow \big(\mathrm{card}(A) < \mathrm{card}(B) \vee \mathrm{card}(A) = \mathrm{card}(B)\big)\,. \tag{23.528}$$

To prove (23.528) we prove the implications

$$\mathrm{card}(A) \leq \mathrm{card}(B) \Longrightarrow \big(\mathrm{card}(A) < \mathrm{card}(B) \vee \mathrm{card}(A) = \mathrm{card}(B)\big)\,. \tag{23.529}$$

and

$$\big(\mathrm{card}(A) < \mathrm{card}(B) \vee \mathrm{card}(A) = \mathrm{card}(B)\big) \Longrightarrow \mathrm{card}(A) \leq \mathrm{card}(B)\,. \tag{23.530}$$

Finally, we have to prove (C5). But (C5) is completely trivial: by definition, "$\mathrm{card}(A) < \mathrm{card}(B)$" means "there is a one-to-one function from A to B but there is no bijection from A to B". So in particular if $\mathrm{card}(A) < \mathrm{card}(B)$ then there is no bijection from A to B, so it's not true that $\mathrm{card}(A) \leq \mathrm{card}(B)$. **Q.E.D.**

To prove the analogues of (R6), (R7), and (R8), we need deeper results from set theory.

Let us start with (C6), the analogue for cardinalities of property (R6). We need to prove that

(&) *if* $\mathrm{card}(A) \leq \mathrm{card}(B)$ *and* $\mathrm{card}(B) \leq \mathrm{card}(A)$ *then* $\mathrm{card}(A) = \mathrm{card}(B)$.

Translating this into English, (&) says:

(&&) *if there exist one-to-one functions* $f : A \to B$, $g : B \to A$, *then there exists a bijection from* A *to* B.

And we have to asnwer the question: ***is (&&true?***.

For finite sets the answer is undoubtedly "'yes": if A, B have cardinalities m, n, and there exist one-to-one functions $f : A \to B$, $g : B \to A$, then it follows from the Dirichlet pigeonhole principle that $m \leq n$ and $n \leq m$, , so $m = n$, and this implies that there exists a bijection from A to B.

Actually, even more can be proved:

Theorem 126. *If A, B are finite sets, and $f : A \to B$, $g : B \to A$ are one-to-one functions, then f is a bijection from A to B and g is a bijection from B to A.*

Proof. The range $\mathrm{Ran}(f)$ of f is a subset of A, and f is a bijection from A to $\mathrm{Ran}(f)$.

Assume that $\mathrm{Ran}(f)$ is a proper subset of B, then $\mathrm{card}(\mathrm{Ran}(f)) < \mathrm{card}(B)$ by Theorem 107, while on the other hand $\mathrm{card}(\mathrm{Ran}(f)) = \mathrm{card}(A)$, so $\mathrm{card}(A) < \mathrm{card}(B)$, and then the Dirichlet pigeonhole principle tells us that there cannot exist a one-to-one function from B to A. Since g is a one-to-one function from B to A, the assumptiom that $\mathrm{Ran}(f)$ is a proper subset of B has led us to a contradiction.

So $\mathrm{Ran}(f) = B$, and then f is onto B. A similar argument proves that g is onto A. **Q.E.D.**

For infinite sets, the analogue of Theorem 126 is not true. Here is a sumple example. Let $\mathcal{E}$ be the set of all even natural numbers. Define $f : \mathbb{N} \to \mathcal{E}$, and $g : \mathcal{E} \to \mathbb{N}$ by letting

$$\begin{aligned} f(n) &= 4n \qquad \text{for} \quad n \in \mathbb{N}\,, \\ g(n) &= n \qquad \text{for} \quad n \in \mathcal{E}\,. \end{aligned}$$

Then $f : \mathbb{N} \to \mathcal{E}$, and $g : \mathcal{E} \to \mathbb{N}$ aare one-to-one functions but neither one is a bijection.

And yet, even though f and g are not themselves bijecyion, a bijection from $\mathbb{N}$ to $\mathcal{E}$ does exist: just define $h : \mathbb{N} \to \mathcal{E}$ by letting $h(n) = 2n$, and it is clear that h is a bijection from $\mathbb{N}$ to $\mathcal{E}$.

It turns out that what happened in this example actually does happen in general.

Theorem 127. *(Cantor-Schroeder-Bernstein) If A, B are sets and and $f : A \to B$, $g : B \to A$ are one-to-one fucntions, then there exists a bijection from A to B.*

The proof of this theorem is given in the book, and I am not going to do it here.

Thanks to the Cantor-Schroeder-Bernstein theorem, statement (&&) above is true, and this implies that condition (C6), the analogue for sets of property (R6), is true. So we get the following theorem:

Theorem 128. *Let A, B be sets such that* $\text{card}(A) < \text{card}(B)$. *Then* $\sim \text{card}(A) = \text{card}(B)$, *that is, B does not have the same cardinality as A.*

(*We do not give a proof, because we have already proved this.*

We now turn to (C7), the cardinality analogue of (R7).

Theorem 129. *If A, B, C are sets such that* $\text{card}(A) \leq \text{card}(B)$, $\text{card}(B) \leq \text{card}(C)$, *and one of the inequalities is strict (that is, either* $\text{card}(A) < \text{card}(B)$ *or* $\text{card}(B) < \text{card}(C)$*) then* $\text{card}(A) < \text{card}(C)$.

Proof. Assume that $\text{card}(A) \leq \text{card}(B)$ and $\text{card}(B) \leq \text{card}(C)$. Then we pick one-to-one functions $f : A \to B$ and $g : B \to C$.

We want to prove that if one of the inequalities is strict, then $\text{card}(A) < \text{card}(C)$.

We already know from Theorem 125, condition (C3), that $\text{card}(A) \leq \text{card}(C)$, and this means that either $\text{card}(A) < \text{card}(C)$ or $\text{card}(A) = \text{card}(C)$.

So, in order to prove that $\text{card}(A) < \text{card}(C)$, we have to exclude the possibility that $\text{card}(A) = \text{card}(C)$.

Suppose that $\text{card}(A) = \text{card}(C)$. Then there exist bijections h, k, from A to C and from C to A, respectively.

So we have functions

$$f : A \to B\,,\ \ g : B \to C\,,\ \ h : A \to C\,,\ \ k : C \to A\,,$$

such that f and g are one-to-one and h and k are bijections.

Since the composite of two one-to-one functions is one-to-one, the composite function $k \circ g : B \to A$ is one-to-one. So we have one-to-one functions $f : A \to B$, $k \circ g : B \to A$. Then the Cantor-Schroeder-Bernstein theorem tells us that there exists a bijection from A to B so $\boxed{\text{card}(A) = \text{card}(B)}$.

Similarly, the composite function $f \circ k : C \to B$ is one-to-one. So we have one-to-one functions $g : B \to C$, $f \circ k : C \to B$. Then the Cantor-Schroeder-Bernstein theorem tells us that there exists a bijection from B to C so $\boxed{\text{card}(B) = \text{card}(C)}$.

So we have proved that both $\text{card}(A) = \text{card}(B)$ and $\text{card}(B) = \text{card}(C)$. And this contradicts the assumption either $\text{card}(A) < \text{card}(B)$ or $\text{card}(B) < \text{card}(C)$.

This contradction arose from assuming that $\text{card}(A) = \text{card}(C)$. Hence it is not true that $\text{card}(A) = \text{card}(C)$. **Q.E.D.**

Finally, we need the cardinality analogue of (R8). This is indeed true, and we state it as a theorem:

Theorem 130. *If A, B are sets, then* $\text{card}(A) \leq \text{card}(B)$ *or* $\text{card}(B) \leq \text{card}(A)$. *That is, one of the following is true:*

- *there exists a one-to-one function from A to B,*
- *there exists a one-to-one function from B to A.*

The proof of this theorem requires methods that are above the level of this course. So here we just leave the theorem without proof.

23.3.5 Infinitely many infinite cardinals

It follows from Theorem 124 that

Theorem 131 *The power set $\mathcal{P}(\mathbb{N})$ is not countable.*

So there are are least two different infinite cardinals: that of $\mathbb{N}$ (and all countably infinits sets), and that of $\mathcal{P}(\mathbb{N})$. But then it is clear that there are many more, in fact infinitely many more, because we can consider the sets

$$\mathbb{N}\,,\ \mathcal{P}(\mathbb{N})\,,\ \mathcal{P}\big(\mathcal{P}(\mathbb{N})\big)\,,\ \mathcal{P}\Big(\mathcal{P}\big(\mathcal{P}(\mathbb{N})\big)\Big)\,,\ \mathcal{P}\left(\mathcal{P}\Big(\mathcal{P}\big(\mathcal{P}(\mathbb{N})\big)\Big)\right),\cdots.$$

That is, we can define, inductively,

$$\begin{aligned}\mathcal{P}_0(\mathbb{N}) &= \mathbb{N}\,,\\ \mathcal{P}_{n+1}(\mathbb{N}) &= \mathcal{P}\Big(\mathcal{P}_n(\mathbb{N})\Big) \ \text{ for } \ n \in \mathbb{Z}\,,\, n \geq 0\,,\end{aligned}$$

and in this way we obtain an infinite sequence $\Big(\mathcal{P}_n(\mathbb{N})\Big)_{n=1}^{\infty}$ of sets each one of which has cardinality strictly larger than the previous obe.

But the story does not end there. We can construct an enormous set $\mathcal{P}_\infty(\mathbb{N})$, defined by

$$\mathcal{P}_\infty(\mathbb{N}) = \bigcup_{n=0}^{\infty} \mathcal{P}_n(\mathbb{N})\,,$$

and then stasrt again, contructiong the sets $\mathcal{P}_n\Bigg(\big(\mathcal{P}_\infty(\mathbb{N})\big)\Bigg)$, and so on.

The result is an infinite tower of infinite towers of infinite towers ,,,, of infinite cardinals.

24 The paradoxes of set theory: Russell's paradox and others

But in the story I have been telling you there are serious problems. We said that the cardinality of the power set of a set X is strictly larger than the cardinality of X. But ***what if we apply this to the set of all sets?***

That is, let

$$X = \Big\{\, x : x \text{ is a set} \,\Big\}.$$

Then we have seen that $\boxed{\text{card}\big(\mathcal{P}(X)\big) > \text{card}(X)}$.

But $\mathcal{P}(X)$ is a set whose members are sets, so $\mathcal{P}(X)$ is a subset of X, and then $\boxed{\text{card}\big(\mathcal{P}(X)\big) \leq \text{card}(X)}$.

So we get a contradiction! And this time we have done nothing wrong!

24.0.6 The Russell paradox

Here is another way to get a contradiction in Set Theory. It is the famous ***Russell paradox.***

Let X be the set of all sets that are not members of themselves. That is, we let

$$X = \{x : x \text{ is a set} \wedge x \notin x\}\,.$$

Let us prove that $X \in X$. Suppose that $\boxed{X \notin X}$. Then is a set that is not a member of itself. So X satisfies the membership criterion for X. Hence $X \in X$. But we are assuming that $X \notin X$. So $\boxed{X \in X \wedge X \notin X}$, which is a contradiction.

So the assumption that $X \notin X$ has led us to a contradiction. Hence $\boxed{\boxed{X \in X}}$.

Now let us prove that $X \notin X$. Suppose that $\boxed{X \in X}$. Then is a set that is a member of itself. So X does not satisfy the membership criterion for X. Hence $X \notin X$. But we are assuming that $X \in X$. So $\boxed{X \in X \wedge X \notin X}$, which is a contradiction.

So the assumption that $X \in X$ has led us to a contradiction. Hence $\boxed{\boxed{X \notin X}}$.

So we have proved that $X \in X \wedge X \notin X$, which is a contradiction.

So

$$\boxed{\boxed{\boxed{\text{In set theory it is possible to prove a contradiction.}}}}$$

And, once you have proved a contradiction, then everything csn be proved, because:

$$\boxed{\boxed{\boxed{\text{In a theory in which it is possible to prove a contradiction, it is possible to prove every statement, whether it is true or false.}}}}$$

This is so for a simple reason: take any statement S you want (for example, S could be "$2 + 2 = 5$", or "6 is a prime number", or "the 45th president of the U.S. is Hillary Clinton", or "the Earth hs flat and rests on top of a giant turtle".

Let us prove S, assuming we know how to prove a contradction C.

To prove S, you do it by contradiction: start with "Assume $\sim S$". Then insert your proof of C, ending with C. So, you see, assuming $\sim C$ we got a contradiction. Therefore we have proved S. **Q.E.D.**

What is wrong with this. Isn't is wonderful that we can prove everything? Shouldn't we celebrate? No more having to study hard to learn how to prove things!!! We have an easy method for proving everything, without doing any work!

The trouble is,

A theory in which it is possible to prove everything is completely useless.

The purpose of writing proofs is to make sure that the mathematical statements we write are true. If we prove something, then we can be sure it is true, because the rules of logic are designed to guarantee that everything we prove is true.

The catch is: if we can prove everything, including statements that are false, then the fact that we have proved something tells us nothing: it could be true or it could be false.

Think of theorem-proving as analogous to smoke-detecting.

A naïve person might think that, since the purpose of a smoke detector is to detect smoke, a device that rings every time there is smoke is precisely what is desired of a good smoke detector.

However, it may happen that the device rigns every time there is smoke because it rings all the time,whetehr there is smoke or not.

Such a "smoke detector" is useless. What you want is a device that rings when there is smoke and does not rignt when there is no smoke. That way, when you hear the device ring, you know that there is smoke.

Similarly, theorem-proving is useful if you can prove the statements that are true, and you cannot prove those that are not true. If you can prove too muc=h, if you can prove assertion that are false, then your theorem-proving has absolutely no value.

24.0.7 The need for Axiomatic Set Theory

The trouble with the paradoxes of the last section arose from the fact that we made indiscriminate use of the Axiom of Set Formation. The axiom says that "if we take any one-variable predicate $P(x)$, we can form the set

$\{x : P(x)\}$." Using this, we created the sets $\{x : x \text{ is a set}\}$ (the set of all sets) and $\{x : x \text{ is a set} \wedge x \notin x\}$ (the set of all sets that do not belong to themselves). And these sets led us to contradictions.

The solution that mathematicians have adopted is to develop "Axiomatic Set Theory" (AST). In AST, axioms are stated that tell us under what conditions is is possible to create a set. And the axioms are carefully chosen so that the sets that caused us trouble cannot be formed.

But this should be the subject of another course.

25 Some more problems

In the following problems, if a and b are real numbers, we write "$[a, b]$" to denote the closed interval $\{x \in \mathbb{R} : a \leq x \leq b\}$, and "$(a, b)$" to denote the open interval $\{x \in \mathbb{R} : a < x < b\}$.

Problem 139. ***Prove*** that if X is a set then there does not exist a one-to-one function $f : \mathcal{P}(X) \longrightarrow X$. □

Problem 140.

1. Let f, g be the functions defined by

   (i) $\text{Dom}(f) = \mathbb{R}$,

   (ii) $f(x) = \frac{x}{\sqrt{1+x^2}}$ for $x \in \mathbb{R}$,

   (iii) $\text{Dom}(g) = (-1, 1)$,

   iv) $g(y) = \frac{y}{\sqrt{1-y^2}}$ for $y \in (-1, 1)$,

   (a) ***Prove*** that $f : \mathbb{R} \longrightarrow (-1, 1)$, $g : (-1, 1) \longrightarrow \mathbb{R}$, $g \circ f = I_{\mathbb{R}}$, and $f \circ g = I_{(-1,1)}$.

   (b) ***Conclude*** from this that $\mathbb{R}$ and the open interval $(-1, 1)$ have the same cardinality.

2. If $a, b \in \mathbb{R}$ and $a < b$, let $f_{a,b}$ be the function with domain $(-1, 1)$, given by $f_{a,b}(x) = a + \frac{1}{2}(b - a)(x + 1)$ for $x \in (-1, 1)$. ***Prove*** that f is a bijection from $(-1, 1)$ to (a, b), and ***conclude*** from this that $\mathbb{R}$ and the interval (a, b) have the same cardinality.

Problem 141. ***Construct*** a bijection from the closed interval $[0, 1]$ to the open interval $(0, 1)$. (Recall that $[0, 1]$ is the set $\{x \in \mathbb{R} : 0 \leq x \leq 1\}$, and $(0, 1)$ is the set $\{x \in \mathbb{R} : 0 < x < 1\}$.)

HINT: Imagine an infinite hotel[123] H, in which

- the points of the open interval $(-1, 1)$ are the numbers of the rooms of H; each room has a number which is a member of $(0, 1)$, and for eacm member x of $(-1, 1)$ there is a room no. x. (So for example there is room no. 0.00001, room no. 0.3, room no. $\frac{1}{\pi}$, room no. $\frac{23}{77}$, room no. 0.9, room no. 0.99, room no. 0.999, etc. But of course there is no room 0 or room 1, because the rooms of H correspond to the members of the *open* interval $(0, 1)$, so 0 and 1 are not possible room numbers.)
- At some time, the hotel has a set of guests, also labeled by the members of the open interval $(0, 1)$, and each guest occupies the room with the same label (so guest no. 0.13 occupies room no. 0.13, guest no. $\frac{1}{\pi}$ occupies room no. $\frac{1}{\pi}$, and so on).
- Suddenly, two new guests, labeled 0 and 1, arrive and ask for rooms.
- And nobody, neither the old guests nor the new ones, is willing to share a room.
- So you have to accomodate these two new guests, while making sure that none of the old guests is left without a room.

If we were dealing with a finite hotel, this would be impossible. If a hotel has 100 rooms, all of which are occupied, and two new guests arrive and ask for a room, there is no way to oblige them.

But for an infinoite hotel it can be done. The way you can do this is by finding within $(0, 1)$ an infinite sequence of rooms (for example, rooms $\frac{1}{2}$, $\frac{1}{3}$, $\frac{1}{4}$, $\frac{1}{5}$, $\frac{1}{6}$, and so on), and then move the guests in those rooms to other rooms also in the sequence, making room for the two new guests.

[123] This ins known as "Hilbert's holtel".