Introduction to Mathematical Proofs*

Attila Máté
Brooklyn College of the City University of New York

March 12, 2020

Contents

*Written for the course Mathematics 2001 (Transition to Advanced Mathematics) at Brooklyn College of CUNY.

1 Introduction

These notes are primarily about proofs, and not the mathematical subjects discussed. The first question about proofs that arises immediately is, can I learn how to do (find, create) proofs. The answer to this question is a clear no. Sometimes centuries elapse before a proof eventually found. Aristotle already claimed that the diameter and the circumference of the circle are incommensurable,[1.1] In other words, he conjectured (guessed) that π is irrational. Yet the first proof of this was only given by Lambert in 1761. Another famous example is Fermat's Last Theorem,[1.2] which, after being first stated without a proof, had to wait more than 350 years for a proof. Many proofs are much easier to find, but finding them is still a challenge. Often one needs to have ideas as to how to proceed that do not seem to have anything to do with the assertion to be proved.

1.1 Reading proofs

Reading proofs is usually much easier than finding then. In reading a proof, all you need to to is to check the correctness of each of the steps performed. This may still be difficult, primarily because of the technical knowledge required, but often also because of the lack of details provided by the author: in a proof, steps are often omitted with the expectation that the reader can fill in the details.

The other problem with reading proofs is often an unusual step occurring in the proof, which may one wonder why such a step is taken. In such a situation, the best approach is to ignore the reason the step is taken, and focus only on the question whether the step is correct. After reading the proof, the reader may return the the question of why; it should be easier to find the reason after one seen the whole proof, but at times the answer is not found even then. Even when someone explains the proof and she knows the reason the step is taken, on first telling the proof might be the best approach to ignore explaining the reason, since it distracts from understanding the logic steps; it is a difficult question how to best explain a proof.

The reason behind a strange step is often not known even to the author. The step may have taken in a sudden flash of inspiration, or perhaps after years of trial and errof, and by the time the proof is finished the reason might be forgotten. The author, with hindsight, may still attribute motivation to the steps taken, even though this motivation may not have anything to do with the original reason. In a different situation, one first finds a complicated proof, and later simplifies it, but then in the simplified proof the reason for the steps can no longer seen. The author, when describing the proof, may fill in some motivation, but that is difficult to do, since then one may have to return to the original, more complicated form of the argument, which does not necessarily help the reader. Often, with proofs written down hundreds of years ago, one will never be able to understand the original motivation behind the proof.

1.2 Learning proofs

Sometimes proofs should be completely memorized, since they contain types of arguments that can be reused in different situtaion. The memorization should definitely not be word for word; it

[1.1] See Definition 15.1 for the definition.
[1.2] See Theorem 6.1.

should be done by remembering the logic steps taken, and perhaps the reasons behind the steps if understood, even if these reasons are *post facto*[1.3] rationalizations, and not the original reasons. In proofs learned years ago, or often even recently, one may not remember the original notation used, yet, by remembering the logical steps, one can reconstruct a proof, usually using a different notation.

One way of learning a proof is, after reading it a small number of times, perhaps only once, one tries to retell the proof from memory, or perhaps write it down; occasially, one gets stuck, and then taking a glimpse at the given proof, one can continue. Then repeating this process until one can reconstruct the whole proof without taking a peek. Using the same words or the same notation, or even the same order of steps as long as the reconstructed proof is correct, is unimportant.

2 Interchange of quantifiers

There are two kinds of quantifiers: universal: $\forall$, meaning "for all," and existential: $\exists$, meaning " there is" or "there exists." Within a quantifier, one may specify the kind of things the quantifier talks about, e.g.,

$$(\forall x : x \text{ is an integer}), \quad (\exists y : y \text{ is a real number}), \quad (\forall x \in \mathbb{Z}), \quad (\exists y \in \mathbb{R}),$$

etc.; here $\mathbb{Z}$ stands for the set of integers (positive, negative, or zero), and $\mathbb{R}$ stands for the set of reals. Two quantifiers of the same kinds mean either two universal quantifiers or two existential quantifiers, while two quantifiers of different kinds refer to one universal and one existential quantifier. Two quantifiers of the same kind are always interchangeable, but two quantifiers of different kinds are not. To see this, consider the following example:

$$(\forall x : x \text{ is licensed driver})(\exists y : y \text{ is a car}) \ (x \text{ has driven } y).$$

This sentence is entirely reasonable, since it only says that "every licensed driver has driven a car," or, more precisely, "every licensed driver has driven at least one car." In real life, there are exceptions even to such a reasonable statements, but if you restrict your attention to life in a small town, the statement is most likely true. If one interchanges the quantifiers, the result is totally absurd:

$$(\exists y : y \text{ is a car}) \ (\forall x : x \text{ is licensed driver}) \ (x \text{ has driven } y).$$

This says that "there is a single car that every licensed driver has driven," and this is unlikely to be true even in a small town, unless everyone has gone to the same driving school within the last few years, and got a chance of practising on the same car.

Another, more mathematical, example is the following. Let x and y run over integers. Then the formula

$$(\forall x)(\exists y)[x > y]$$

is true, while the formula

$$(\exists y)(\forall x)[x > y]$$

is false. Indeed, the first formula is true. Given an arbitrary integer x, we can pick $y = x - 1$ to ensure that $x > y$. On the other hand, the second formula is not true. To see this, pick an arbitrary number y. Then $(\forall x)[x > y]$ is certainly not true; for example, $x > y$ is not true with $x = y - 1$.

[1.3] Latin for "after the fact."

3 More on logic

A sentence is a statement that is either true or false.[3.1] Sentences can be connected by the logic operations, also called sentential connectives *and*, *or*, *if ... then*, and *if and only if*, denoted in turn by $\&$, $\vee$, $\rightarrow$, and $\leftrightarrow$. Instead of $\&$, people often use the symbol $\wedge$. The logic operations only connect sentences (or, more generally, statements – see below). For example, the "or" in the sentence "You can have coffee or tea with your breakfast" is not a logic operation.[3.2]

The meaning of the sentential connectives can be illustrated by truth tables. Writing T for true and F for false (T and F are called *truth values*) we have for the operation $\&$, called *conjunction*:

A	B	$A \,\&\, B$
T	T	T
T	F	F
F	T	F
F	F	F

For the operation $\vee$, called *disjunction*, we have

A	B	$A \vee B$
T	T	T
T	F	T
F	T	T
F	F	F

As seen from the truth table, disjunction is always meant in the *inclusive* sense, that is $A \vee B$ is true unless both A and B are false. This differs from colloquial usage, where one often uses the word "or" in the exclusive sense, where "A or B in the exclusive sense" is true if exactly one of A and B is true.

For the operation $\rightarrow$, called *conditional*, we have

A	B	$A \rightarrow B$
T	T	T
T	F	F
F	T	T
F	F	T

The meaning of the conditional is often differs from its colloquial use, where the meaning of "if A then B" is unclear in case A is false. In mathematics, one strictly follows the truth table above. For example, the sentence "if 2 by 2 is 5 then the snow is black" is a true sentence in mathematics.[3.3] In

[3.1] That is, true or false in principle. If one talks about formal logic, then a sentence would be a declarative sentence. Some declarative sentences, however, cannot be considered sentences in the sense of logic, since they cannot be considered true or false even in principle. For example, the sentence "This sentence is false" cannot be taken as true (since it then says that it is false), and it cannot be taken as false (since then it says that it is true). In mathematical logic, what is a sentence or a statement is defined formally by describing the rules how a statement can be formed. The collection of these rules is called *syntax*.

[3.2] However, on can rewrite this as a logic operation by saying that "You can have coffee with your breakfast or you can have tea with your sentence" (although by doing this, one probably changes the meaning, since in logic, "or" is meant in the inclusive sense (allowing to have both coffee and tea), while on a restaurant menu the meaning is probably exclusive (not allowing to have both coffee and tea without paying extra).

[3.3] This example is due to the German mathematician David Hilbert. Note that the sentence "if 2 by 2 is 5 then the snow is white" is also true.

colloquial speech one would consider this sentence meaningless, or at best pointless. But it illustrates an important point: in the conditional, there does not need to be a causal connection between the constituents.[3.4] For the operation $A \to B$, instead of saying "if A then B", it is often more convenient to say "A only if B". In case of this latter sentence, the colloquial meaning approaches more closely the mathematical meaning of the conditional. Namely, "A only B" means that A is allowed to be true only in case B is also true; indeed, when A is true and B is false, the truth table entry for $A \to B$ shows false.

One occasionally reverses the arrow in the conditional, using the symbol $A \leftarrow B$ meaning "A if B", or "if B then A":

A	B	$A \leftarrow B$
T	T	T
T	F	T
F	T	F
F	F	T

The conditional $A \leftarrow B$ (or $B \to A$) is usually called the *converse* of $A \to B$. Finally, the truth table of the operation $\leftrightarrow$, called *biconditional*, is

A	B	$A \leftrightarrow B$
T	T	T
T	F	F
F	T	F
F	F	T

$A \leftrightarrow B$ is expressed as "A if and only if B", or, sometimes, as 'A iff B" (but it is not clear how one should pronounce the word "iff"). The word "iff" was introduced by Paul Halmos. "A if and only if B" is short for saying "A if B and A only if B"; formally, for $(A \leftarrow B)\,\&(A \to B)$. It is easy to check that the truth table for this formula is the same as the one given for the biconditional above.

These logic operations are called *binary* operations, since they involve two constituents, called *operands*, A and B in the above truth tables. The letters A and B used in the above formulas are often called *sentential variables*, i.e., variables that can either be true or false.

3.1 Negation

operation "not" is called *negation*. "Not A means that "it is not the case that A", or, more simply, "A is not true". This is called a *unary* operation, since it has only one operand. The truth table for negation is

A	$\neg A$
T	F
F	T

3.2 Tautologies

tautology is a logic expression (an expression involving the logical operations just defined) that is always true, whether or not the sentential variables in it are true or false. Examples for tautologies are

$$A \vee \neg A, \quad A \to A, \quad \neg\big(A\,\&(\neg A)\big).$$

[3.4]That is, 2 by 2 being 5 does not cause the snow to be black. As for the constituents, or, with a more technical word, *operands*, in the conditional $A \to B$, A called the antecedent and B is called the consequent.

3.2.1 Tautologies showing the equivalence of two formulas

Some more interesting tautologies have form

$$\text{formula}_1 \leftrightarrow \text{formula}_2.$$

Such tautologies say that formula_1 and formula_2 say the same thing, hence formula_1 can be replaced by formula_2. For example, we have

$$(\neg(A \,\&\, B)) \leftrightarrow ((\neg A) \vee (\neg B)).$$

For better readability, one can drop several pairs of parentheses here, to write

$$\neg(A \,\&\, B) \leftrightarrow \neg A \vee \neg B.$$

To make sense of this way of writing the formula, one can assigns priority to the logic operations in the order $\neg$, $\&$, $\vee$, $\leftarrow$, $\rightarrow$, (k+1)* $\leftrightarrow$, meaning that one first try to perform the operations with higher priority.[3.5] Most people consider the priority between $\&$ and $\vee$, and between $\leftarrow$ and $\rightarrow$ unclear, so it is best to use parentheses to avoid misunderstanding. To check that the above formula is indeed a tautology, one can use the truth table to evaluate it for each choice of A and B to find that the formula is always true. One interpretation of the above tautology is that $\neg(A \,\&\, B)$ and $\neg A \vee \neg B$ mean the same thing. This is because the biconditional is true exactly when the two sides have the same truth value; so the above formula being always true means that the two sides on the biconditional in it always have the same truth value. The above formula is one of the two *De Morgan identities*. The other De Morgan identity is the tautology

$$\neg(A \vee B) \leftrightarrow \neg A \,\&\, \neg B. \tag{3.1}$$

Another simple tautology is

$$(A \rightarrow B) \leftrightarrow \neg A \vee B; \tag{3.2}$$

we used parentheses on the left here, since the sometimes $\rightarrow$ and $\leftrightarrow$ is considered to have equal priority. An even simpler tautology is

$$\neg\neg A \leftrightarrow A. \tag{3.3}$$

This is also called the law of excluded middle; this law is expressed in the Latin phrase "tertium non datur."[3.6] When one wants to establish the *implication*[3.7] $A \rightarrow B$, one often takes advantage of the tautology

$$(A \rightarrow B) \leftrightarrow (\neg B \rightarrow \neg A), \tag{3.4}$$

[3.5] This is similar to the rule in algebra that $\cdot$ (multiplication) has higher priority than $+$ (addition). That is, the formula $2+3\cdot 5$ means $2+(3\cdot 5)$, and not $(2+3)\cdot 5$. $+$ and $-$ have equal priority, so in the expression $2+3-5-6+8$ one performs the operations from left to write, i.e., this expression means $(((2+3)-5)-6)+8$ In computer science, one often uses the word *precedence* instead of priority; computer scientists unambiguously assign higher precedence to to logical and than to logical or; in mathematics, this precedence is not always taken for granted.

[3.6] The phrase means "There is no third [possibility]." This tautology is not accepted in intuitionistic mathematics, an approach to the foundation of mathematics that was precipitated by the crisis in set theory at the end of the 19th century. Mainstream mathematics does not accept the intuitionistic approach; instead, it prefers an approach via axiomatic set theory.

[3.7] One often calls a conditional an implication, especially in informal usage, since instead of "if A then B" one might say "A implies B". However, logical implication has a meaning separate from, though related to, that of the conditional. Therefore, some consider calling a conditional an implication objectionable.

and proves the implication $\neg B \to \neg A$ instead. The conditional $\neg B \to \neg A$ is called the *contrapositive* of $A \to B$. A further tautology is

$$\neg(A \leftrightarrow B) \leftrightarrow (\neg A \leftrightarrow B).$$

One often writes $A \nleftrightarrow B$ instead of $\neg(A \leftrightarrow B)$. The truth table of $A \nleftrightarrow B$ is

A	B	$A \nleftrightarrow B$
T	T	F
T	F	T
F	T	T
F	F	F

One might call the operation $A \nleftrightarrow B$ *exclusive or*, since it reflects the meaning the word "or" often used colloquially. However, it is best to avoid this term "exclusive or", since in mathematics the word "or" is always used in the inclusive sense, as defined by the logic operation $A \vee B$. The expression $A \nleftrightarrow B$ in mathematics is often given as "A if and only if not B", reflecting the fact that this this expression is equivalent to (i.e., true exactly the same time as) the expression $A \leftrightarrow \neg B$; that is, the fact that

$$(A \nleftrightarrow B) \leftrightarrow (A \leftrightarrow \neg B)$$

is a tautology.

3.3 Equivalence versus biconditional

The equivalence relation $\equiv$ asserts that the logic expessions on the two sides have the same truth value. For example $A \equiv B$ says that either both A and B are true, or they are both false. $A \leftrightarrow B$ appears to say the same thing in that it is true if both A and B are true, or if both A and B are false, and false in other cases. But there is a difference: $\equiv$ is a relation symbol, and $\leftrightarrow$ is a binary logic operation. The difference can be appreciated by comparing equality $=$, a relation symbol, and addition $+$, an operation. We can say that $5 = 3+2 = 6$, and one would say that the first equality is true, the second one is false, but one would not assign truth or falsity of the whole of $5 = 3 + 2 = 6$. On the other hand the formula $(4 + 2) + 3$ produces the single number 9. The fact that addition is associative, and so $4 + (2 + 3)$ produces the same number is incidental. But because of associativity, one omits the parentheses, and writes $4 + 2 + 3$; this, however, does not change the fact that one computes the result by adding two numbers and then one adds a third one.[3.8]

The situation is similar in logic. $A \equiv B \equiv C$ constitutes two separate assertions: A has the same truth value as B, and B has the same truth value as C. On the other hand, to interpret the formula $A \leftrightarrow B \leftrightarrow C$, and needs to use parentheses, either as $(A \leftrightarrow B) \leftrightarrow C$ or $A \leftrightarrow (B \leftrightarrow C)$. There are no agreed rules as to which one is the correct reading, so dropping the parentheses may be undesirable. Fortunately, both placement of parentheses procudes the same truth value; indeed the operation is associative, as we will explain. However, the result is unexpected, and not the one we wanted to

[3.8] Even when you add a whole column of numbers on paper, you go down the columns adding digits one by one, rather than figuring out the sum of the whole column just by looking at it.

express. Here is the truth table of the first of these formulas:

A	B	C	$A \leftrightarrow B$	$(A \leftrightarrow B) \leftrightarrow C$
T	T	T	T	T
T	T	F	T	F
T	F	T	F	F
T	F	F	F	T
F	T	T	F	F
F	T	F	F	T
F	F	T	T	T
F	F	F	T	F

By perusing the truth table, one can ascertain that $(A \leftrightarrow B) \leftrightarrow C$ is true if exactly one or three among A, B, C is true. The same is true for the formula $A \leftrightarrow (B \leftrightarrow C) \equiv (B \leftrightarrow C) \leftrightarrow A$, so, indeed we have $(A \leftrightarrow B) \leftrightarrow C \equiv A \leftrightarrow (B \leftrightarrow C)$, as we claimed above. It is, however, also clear from this formula that it does not say the same thing as $A \equiv B \equiv C$.

3.4 Open statements

An *open statement* is a statement that has (zero or more) variables in it; when one gives values to these variables. For example, to say that "x is greater than 2" (where x denotes an unspecified real number) is on open statement. One can tell its truth value only after one specifies what real number x is. In mathematical logic, one describes the rules how open statements are formed; the collection of these rules are called *syntax*. However, want to discuss matters somewhat informally, so we will avoid a detailed discussion of syntax. Occasionally, open statements will be denoted by script capital letters such as $\mathcal{A}$ or $\mathcal{B}$.

3.5 Terms, free, and bound variables

A variable that is not quantified is called a *free* variable. A variable that is quantified is called a *bound* variable. An expression that designates an object in the system is called a *term*. For example, when one studies integers, an expression such as

$$(x+3)^2 + 5$$

is a term. A variable itself is also a term.

3.5.1 Substitutability

If ϕ is a formula, x is a variable, and t is a term, $\phi(t/x)$ denotes the formula obtained by substituting (i.e., replacing) all free occurrences of x with t. In a less formal context, one may write $\phi(x)$ to indicate that x may (or may not) occur in the formula $\phi(x)$,[3.9] and instead of $\phi(x)(t/x)$ one would write $\phi(t)$.

In any case, one expects (mistakenly), that if $\forall x \phi$ is true then $\phi(t/x)$ is true; after all, if something is true for all x, then it should be true for a specific choice of x. However, one needs to specify the a condition called substitutability.

[3.9] If one write $\phi(x)$, one should not write ϕ instead.

Definition 3.1 (Substitutability)**.** The term t is *substitutable* or *free*[3.10] for x in the formula ϕ if no free variable occurring in t becomes bound by replacing all free occurrences of x by t.

The formula $\forall x \exists y(y > x)$, interpreted over all integers, is true. After all, given any integer x, there is a larger integer. The term $t = y + 1$ is, however, not substitutable for x in the formula $\exists y(y > x)$, since the substitution results in the formula $\exists y(y > y + 1)$, which is obviously false. What happens here is that a free variable in t becomes captured by a quantifier when we perform the substitution.

3.6 Negating quantified statements

The formula $\neg(\forall x)\mathcal{A}$ is the same as $(\exists x)\neg\mathcal{A}$; here $\mathcal{A}$ is some open statement.[3.11] That is, saying that "it is not true that for all x $\mathcal{A}$ holds" means that "there is an x for which A does not hold". One writes this by saying that

$$\neg(\forall x) \equiv (\exists x)\neg;$$

here $\equiv$ means that what are written on the two sides are equivalent, i.e., they mean the same thing, i.e. they can replace each other.[3.12] Similarly, the formula $\neg(\exists x)\mathcal{A}$ is the same as $(\forall x)\neg\mathcal{A}$. That is, to say that "it is not true that there exists an x for which $\mathcal{A}$ holds" means that "for all x it is not true that $\mathcal{A}$ holds". This can be expressed by the rule

$$\neg(\exists x) \equiv (\forall x)\neg.$$

One can use these rules and the tautologies mentioned above to move negation inside a formula. For (3.5) example,

$$\neg(\forall \epsilon > 0)(\exists \delta > 0)(\forall p \in (0,1))(\forall q \in (0,1)) \left(|p - q| < \delta \to \left| \frac{1}{p} - \frac{1}{q} \right| < \epsilon \right). \tag{3.5}$$

Here p, q, ϵ, and δ denote real numbers,[3.13] and $(0,1)$ denotes the interval $\{t : 0 < t < 1\}$. What this formula expresses is irrelevant for our present purpose.[3.14] The quantifiers $(\forall \epsilon > 0)$ is called a *restricted quantifier*, since it says that ϵ ranges over positive real numbers instead of all real numbers (where unrestricted variables range in the present case). Similarly, $(\forall \delta > 0)$, $(\exists p \in (0,1)$ are restricted quantifiers. The important point at present is that the above rules of interchanging negation and quantifiers are also true for restricted quantifiers.

[3.10] Both terms are used.

[3.11] Presumably, $\mathcal{A}$ depends on x, so one might be tempted to write $\mathcal{A}(x)$ instead. However, writing $\mathcal{A}(x)$ really does not make things clearer.

[3.12] We do not regard $\equiv$ as a logical connective here. It can be considered as a rule according to which we can change (or transform) formulas without changing their meanings. That is, it expresses a transformation rule; see below.

[3.13] One also says that these variables *range over* real numbers.

[3.14] It expresses the statement that the function $f(x) = 1/x$ is not uniformly continuous in the interval $(0,1)$, discussed below in these notes.

The above formula can be transformed as follows:[3.15]

$$\begin{aligned}
&\neg(\forall\epsilon>0)(\exists\delta>0)(\forall p\in(0,1))(\forall q\in(0,1))\left(|p-q|<\delta\to\left|\frac{1}{p}-\frac{1}{q}\right|<\epsilon\right)\\
&\quad\equiv(\exists\epsilon>0)\neg(\exists\delta>0)(\forall p\in(0,1))(\forall q\in(0,1))\left(|p-q|<\delta\to\left|\frac{1}{p}-\frac{1}{q}\right|<\epsilon\right)\\
&\quad\equiv(\exists\epsilon>0)(\forall\delta>0)\neg(\forall p\in(0,1))(\forall q\in(0,1))\left(|p-q|<\delta\to\left|\frac{1}{p}-\frac{1}{q}\right|<\epsilon\right)\\
&\quad\equiv(\exists\epsilon>0)(\forall\delta>0)(\exists p\in(0,1))\neg(\forall q\in(0,1))\left(|p-q|<\delta\to\left|\frac{1}{p}-\frac{1}{q}\right|<\epsilon\right)\\
&\quad\equiv(\exists\epsilon>0)(\forall\delta>0)(\exists p\in(0,1))(\exists q\in(0,1))\neg\left(|p-q|<\delta\to\left|\frac{1}{p}-\frac{1}{q}\right|<\epsilon\right)\\
&\quad\equiv(\exists\epsilon>0)(\forall\delta>0)(\exists p\in(0,1))(\exists q\in(0,1))\neg\left(\neg(|p-q|<\delta)\vee\left|\frac{1}{p}-\frac{1}{q}\right|<\epsilon\right)\\
&\quad\equiv(\exists\epsilon>0)(\forall\delta>0)(\exists p\in(0,1))(\exists q\in(0,1))\left(|p-q|<\delta\,\&\,\neg\left|\frac{1}{p}-\frac{1}{q}\right|<\epsilon\right)\\
&\quad\equiv(\exists\epsilon>0)(\forall\delta>0)(\exists p\in(0,1))(\exists q\in(0,1))\left(|p-q|<\delta\,\&\left|\frac{1}{p}-\frac{1}{q}\right|\geq\epsilon\right);
\end{aligned}$$

the third line from below was obtained by using the tautology in equation (3.2), and the second line from below was obtained by using equations (3.1) and (3.3).

While it is not important in the present context, at this point, it is easy to show that the above formula is true. To this end, simply pick $\epsilon=1$, then pick an arbitrary $\delta>0$, then pick a $p\in(0,1)$ with $p<\delta$, then pick $q=p/2$. We have

$$\left|\frac{1}{p}-\frac{1}{q}\right|=\left|\frac{1}{p}-\frac{1}{p/2}\right|=\left|\frac{1}{p}-\frac{2}{p}\right|=\left|-\frac{1}{p}\right|=\frac{1}{p}>1;$$

the last equality and the inequality holds because $p\in(0,1)$. As $\epsilon=1$, this shows that we certainly have

$$\left|\frac{1}{p}-\frac{1}{q}\right|\geq\epsilon.$$

As

$$|p-q|=\left|p-\frac{p}{2}\right|=\frac{p}{2}<\delta,$$

this shows that the above formula is indeed true.

The formula expresses a fact that will be familiar in a course usually taken only after the present course, discussing a rigorous introduction to mathematical analysis (commonly called calculus in undergraduate courses): it says that the function $f(x)=1/x$ is not uniformly continuous on the interval $(0,1)$.

3.7 Restricted quantifiers

If E is a set, and $\mathcal{A}$ is an open statement, quantifiers of the form $(\forall x\in E)$ and $(\exists x\in E)$ are called *restricted quantifiers*. More generally, given an open statement $\mathcal{A}$, one can consider the restricted quantifiers $(\forall x:\mathcal{A}$ and $\exists x:\mathcal{A})$. If $\mathcal{B}$ is another open statement, then the formula $(\forall x:\mathcal{A})\mathcal{B}$ says that "for all x such that $\mathcal{A}$ holds we (also) have $\mathcal{B}$, and for $(\exists x:\mathcal{A})\mathcal{B}$ means that "for all x for which $\mathcal{A}$ holds we also have $\mathcal{B}$". It is easy to see that

$$(\forall x:\mathcal{A})\mathcal{B}\equiv(\forall x)(\mathcal{A}\to\mathcal{B}),\tag{3.6}$$

[3.15] See Subsection 3.3 concerning the use of the symbol $\equiv$.

and

$$(\exists x : \mathcal{A})\mathcal{B} \equiv (\exists x)(\mathcal{A} \,\&\, \mathcal{B}). \tag{3.7}$$

Therefore, restricted quantifiers are not strictly necessary, but they are often convenient to use. They frequently make formulas simpler, and, a very important point, the rules discussed above involving the interchange of negation and quantifiers are also true for restricted quantifiers. Finally, for a set E,

$$(\forall x \in E) \equiv (\forall x : x \in E)$$

and

$$(\exists x \in E) \equiv (\exists x : x \in E).$$

3.8 First order logic

What we discussed above is called *first order logic*. That is, we are given a *domain* D in which the variables are allowed to assume values.[3.16]. This distinguishes it from higher order logic, in which one is allowed to quantify over all subcollections of D.

3.9 Reading

[2, Chapter 2, pp. 38–81].

3.10 Homework

[2, Chapter 2, pp. 38–81] p. 40, 2.1, 2.3, p. 42, 2.11, 2.13, p. 45, 2.15, 2.17, p. 47, 2.19, 2.29, p. 52, 2.31, 2.33, p. 56, 2.39, p. 59, 2.47, 2.49, p. 61, 2.53, 2.57. p. 64, 2.63, p. 73, 2.67, 2.71, 2.75, 2.79, p. 77, 2.85, p. 78, 2.91, 2.101, 2.105.

4 Sets

A set is a collection. The members of this collection are called its elements; the symbol $x \in A$ indicates that x is an element of the set A. We can describe a set A by listing its elements inside braces { and }; for example,

$$A = \{1, 3, 5, 7, 9\} \tag{4.1}$$

is the set whose elements are the integers 1, 2, 3, 5, 7, and 9. Certain sets commonly occurring in mathematics have standard notation; for example $\mathbb{Z}$ is the set of all integers (positive, negative, or zero), $\mathbb{Q}$ is the set of all rationals, and $\mathbb{R}$ is the set of all real numbers. Sets can also be described by the set-builder operation:

$$S = \{x : \phi(x)\}$$

denotes the set of all things x for which the condition $\phi(x)$ is satisfied. Here the variable x usually has a certain range, i.e., it can assume certain values specified in the context (for example, one might

[3.16] We call it a domain, which may be thougth of as a set, except we do not want to exclude a discussion of set theory, in which the domain would be the collection (or *class* of all sets, which, as we will see later, cannot be a set. In set theory, class is a technical word, so we did not want to call D a class

agree that x is a real number, i.e., that x runs over real numbers). Sometimes one can indicate the range in the set-builder operation. For example, the set A in equation (4.1) can also be described as

$$A = \{x \in \mathbb{Z} : 1 \leq x < 10 \,\&\, x \text{ is odd}\}; \tag{4.2}$$

here $\&$ is the logical "and." That is, for two statements Φ and Ψ, the statement $\Phi \,\&\, \Psi$ is true only in case both Φ and Ψ are true. The fact that the sets described in formulas (4.1) and (4.2) are the same is a special case of the following

Axiom 4.1 (Axiom of Extensionality)**.** Two sets are equal if and only if they have the same elements.

4.1 The empty set

Once one uses the set-builder operation, it is almost inevitable that one encounters a set with no elements; such a set is called the *empty set*, denoted as $\emptyset$. By the Axiom of Extensionality (Axiom 4.1, the empty set is unique, that is, there is only one empty set. With the set-builder operation, one might occasionally write $\emptyset = \{x : x \neq x\}$. With the listing notation, sometimes one writes $\emptyset = \{\}$, noting that nothing is listed between the braces.

4.2 Relations between sets

Definition 4.1. Given two sets A and B, we say that A is a *subset* of B if every element of A is an elemnt also of B.

In this case, we also say that B is a *superset* of A also this term is used less often than the term subset. The symbol $A \subset B$ expresses the statement that A is a subset of B. We also say that B includes A. The symbol $B \supset A$ can also be used. The use of the word "contain" should be used with extreme care, since it is often misused. B contains A should properly mean that A is an element of B (yes, a set can be an elememnt of another set), but it is often misused to mean that A is a subset of B. Such a misuse should absolutely avoided, The best way to say that $x \in A$ is that x is an element of A, or that x belongs to A.

To illustrate the difference between $\in$ and $\subset$, note that $\emptyset \notin \emptyset$, since the empty set has no element, while $\emptyset \subset \emptyset$ is vacuously true. In fact, for every set A, we have $\emptyset \subset A$ is satisfied: given that $\emptyset$ has no element, the no requirements are imposed on A by saying that every element of $\emptyset$ is also an element of A.

The set $\{\emptyset\}$ is the set whose only element is $\emptyset$; it differs from $\emptyset$, since the former has one element, the latter has none. The sets $\{\emptyset\}$ and $\{\{\emptyset\}\}$ both have one elements. but they are not the same sets according to Axiom 4.1, since their elemnts are not the same, as we saw just before.

4.3 Set operations

Given sets A and B, their *union* $A \cup B$ is the set that *contains* (correct use!) the elements of either:[4.1] that is,

$$A \cup B \stackrel{def}{=} \{x : x \in A \vee x \in B\},$$

where $\vee$ is the symbol for logical "or." That is, for two statements Φ and Ψ, the statement $\Phi \vee \Psi$ is true if Φ or Ψ is true.[4.2]

[4.1] One might say, "contains the elements of both," but such use is ambiguous; this is why we clarify what we mean next.

[4.2] In mathematics, logical "or" is always meant in the inclusive sense; that is $\Phi \vee \Psi$ is true if one of Φ and Ψ is true, or if both are true.

The *intersection* $A \cap B$ of the sets A and B is the set that contains only the elements that belong to both A and B. That is,

$$A \cap B \stackrel{def}{=} \{x : x \in A \,\&\, x \in B\}.$$

The set difference of A minus B, denoted as $A \setminus B$,[4.3] is defined as the set of those elemments of A that are not elements of B:

$$A \setminus B \stackrel{def}{=} \{x : x \in A \,\&\, x \notin B\}.$$

One might read the right-hand side here as the set of those elements x for which $x \in A$ *but* $x \notin B$. While the word "but" expresses contrast, its meaning in this context is not any different from that of the word "and."

The symmetric difference of the sets A and B is defined as the set

$$A \triangle B \stackrel{def}{=} (A \setminus B) \cup (B \setminus A).$$

If A is a set of sets, i.e., a set all whose elements are also sets then $\bigcup A$ is the union of all elements of x. Formally,

$$\bigcup A \stackrel{def}{=} \{x : \text{ we have } x \in B \text{ for some } B \in A\},$$

or, even more formally,

$$\bigcup A \stackrel{def}{=} \{x : (\exists y)(y \in A \,\&\, x \in y)\},$$

where $(\exists y)$ is an *existential quantifier*, to be read as "there is a y such that ...." Using *restricted quantifiers*, this can also be written as

$$\bigcup A \stackrel{def}{=} \{x : (\exists y \in A)(x \in y)\}.$$

Often, mathematicians not trained in logic, and, for an irrational reason, prefer the notation where the elements of A are indexed. That is, let

$$A = \{B_\iota : \iota \in I\};$$

that is, I is a set indexing the elements of A, and the Greek letter ι (iota) is used to indicate that I may not be a set of integers. Assuming that B_ι is a set for all $\iota \in I$, they prefer to use the symbol

$$\bigcup_{\iota \in I} B_\iota,$$

even though the simpler symbol $\bigcup A$ means the same thing.

Similarly, if A is a set of sets, then $\bigcap A$ can be defined as the intersection of all elements of A:

$$\bigcap A \stackrel{def}{=} \{x : \text{ we have } x \in B \text{ for all } B \in A\},$$

or, even more formally,

$$\bigcap A \stackrel{def}{=} \{x : (\forall y)(y \in A \to x \in y)\};$$

[4.3] Sometime the notation $A - B$ is used, but it should be avoided, since the latter notation is also used with different meanings. In earlier times, typesetting $A \setminus B$ caused extra difficulty for printers, but with computerized typesetting that is no longer an issue.

here $(\forall y)$ is a universal quantifier, to be read as "for all y we have ...," and for two statements Φ and Ψ, $\Phi \to \Psi$ is the conditional, meaning "if Φ is true then Ψ is also true". The only time $\Phi \to \Psi$ is false is in case Φ is true and Ψ is false. Using *restricted quantifiers*, this can also be written as

$$\bigcap A \stackrel{def}{=} \{x : (\forall y \in A)(x \in y)\}. \tag{4.3}$$

One needs to be a little careful with using the symbol $\bigcap A$, since it is meaningless in case A is the empty set, since if x in equation (4.3) runs over all sets (as one would naturally expect), then the right-hand side describes the set of all sets, which is meaningless.[4.4] Even if A is the empty set and C is another set, it is reasonable to interpret $C \cap (\bigcap A)$ to be the set C.

4.4 Venn diagrams

Set operations can be illustrated in Venn diagrams. One draws two or three circles in a box, indicating two or three sets, and then shades the results of various set operations. In Figure 4.1, *a)* illustrates the set $A \cap B \cap C \cap$, which is defined as the set $A \cap (B \cap C) = (A \cap B) \cap C$; the equality here can easily be proved, and justifies the omitting of the parentheses in the first expression. Similarly, *a)* illustrates the set $A \cup B \cup C \cup$, which is defined as the set $A \cup (B \cup C) = (A \cup B) \cup C$; the equality here can easily be proved, and justifies the omitting of the parentheses in the first expression. Finally, *a)* illustrates the set $(A \triangle B) \setminus C$. Venn diagrams may be helpful in illustrations, but they should not be used for proofs.

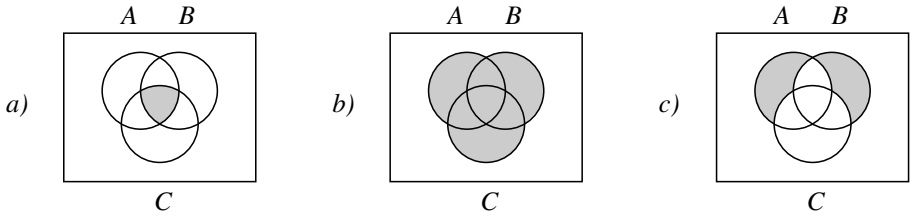

Figure 4.1: Venn diagrams

4.5 The power set of a set

The set of all subsets of a set A is called its power set, and it is usuallt denoed by $\mathcal{P}(A)$. That is,

$$\mathcal{P}(A) \stackrel{def}{=} \{x : x \subset A\}.$$

4.6 Reading

[2, Chapter 1, pp. 14–37].

4.7 Homework

[2, Chapter 1, pp. 14–37], p. 26, 1.29, p. 30, 1.37, 1.43, 1.45, p. 33, 1.49, p. 34, 1.59, 1.63, 1.67, p. 36, 1.89.

[4.4] That is, it is meaningless in most versions of axiomatic theory. It make sense in Quine's New Fundation, and axiomatization of set theory of interest to mathematical logicians and to philosophers, but is not commonly used in mathematical practice. See [23].

5 Prime factorization

More discusion of divisibility will be given in Section 12, where the notation is also explained after Definition 12.1. But we want to discuss the results in this section early, since they are basic in many arguments.

5.1 Two proofs of Euclid's lemma

Lemma 5.1 (Euclid)**.** *Let p be a prime, and let a, b be integers. If $p \mid ab$ then $p \mid a$ or $p \mid b$.*

There are many ways to prove this lemma.

First Proof. Assume p is the smallest prime for which this assertion fails, and let a and b be such that $p \mid ab$ and $p \nmid a$ and $p \nmid b$. By replacing a and b with their remainders when dividing by p, we may assume that $1 \le a < p$ and $1 \le b < p$. Then $kp = ab$; clearly, $1 \le k < p$. We have $k \ne 1$ since p is a prime. Let q be a prime divisor of k. Then $q \mid ab$, and so, by the minimality assumption on p, we have $q \mid a$ or $q \mid b$. Then dividing q into k and into one of a or b, we obtain an equation $k'p = a'b'$, where $1 \le k' < k$, $1 \le a' < p$, and $1 \le b' < p$. Repeating this step as long as necessary, we arrive at an equation $k''p = a''b''$ with $k'' = 1$, $1 \le a'' < p$, and $1 \le b'' < p$. This equation contradicts the primality of p, completing the proof. □

The second proof gives Euclid's Lemma is a corollary of the following.

Lemma 5.2. *Let let a and c be positive integers and let t be the smallest positive integer such that $c \mid at$. Let b be a positive integer such that $c \mid ab$. Then $t \mid b$. In particular, $t \mid c$.*

Proof. Assume $t \nmid b$; let q and r be such that $b = tq + r$ and $1 \le r < t$. Then

$$ab = atq + ar.$$

As c is a divisor of the left-hand side and of the first term on the right-hand side, it follows that c is also a divisor of the second term on the right-hand side; i.e., $c \mid ar$. This, however, contradicts the minimality of t. This contraction shows that $t \mid b$. Since we have $c \mid ac$, this assertion with $c = b$ shows that $t \mid c$ holds. □

Corollary 5.1 (Euclid)**.** *Let p be a prime, and let a and b be positive integers. If $p \mid ab$ then $p \mid a$ or $p \mid b$.*

Proof. Let t be the smallest positive integer such that $p \mid at$. Then we have $t \mid b$ and $t \mid p$ by Lemma 5.2. The latter implies that $t = 1$ or $t = p$. In the former case we have $p \mid a$, in the latter case we have $p \mid b$. □

Corollary 5.2. *Let a, b, and c be positive integers such that $(b, c) = 1$. If $c \mid ab$ then $c \mid a$.*

Proof. Let t be the smallest positive integer such that $c \mid at$. Then we have $t \mid b$ and $t \mid c$ by the Lemma. As $(b, c) = 1$, we must have $t = 1$. Since $c \mid at$, this means that $c \mid a$. □

5.2 Unique prime factorization

Theorem 5.1 (Fundamental Theorem of Arithmetic)**.** *Any integer greater than* 1 *can be written as a product of primes in only one way, except for the order in which the primes are written.*

There are many ways to prove this Theorem. The one we present avoids the use of Euclid's Lemma 5.1.

Proof. Let $n > 1$ be the smallest integer that has two different prime factorizations, and let p be the smallest prime that occurs in any prime factorization of n. The prime p can occur only in one prime factorization of n; indeed, if p occurred in two different prime factorizations of n, writing the equation expressing the equality of these two prime factorizations, we could cancel p from both sides and obtain an integer smaller than n with two different prime factorizations. Since $p \mid n$, we have $n = mp$ for some integer m. Let $q_1 q_2 \dots q_k$ be a prime factorization of n in which p does not occur. We have

$$mp = q_1 q_2 \dots q_k,$$

where $q_i > p$ for each i with $1 \leq i \leq k$. When dividing p into q_i, denote the quotient by l_i, and the remainder by r_i; we have $q_i = l_i p + r_i$ and $0 < r_i < p$. That is

$$mp = (l_1 p + r_1)(l_2 p + r_2) \dots (l_k p + r_k).$$

If we multiply out the right-hand side, all terms will contain p except for the term $r_1 r_2 \dots r_k$. Adding up the terms containing p, we get a multiple of p; write this as ap, where a is an integer. That is

$$mp = ap + r_1 r_2 \dots r_k.$$

Writing $b = m - a$, we have

$$bp = r_1 r_2 \dots r_k.$$

Here $b > 0$, since the right-hand side is positive. Taking the prime factorizations of b and of each r_i for i with $1 \leq i \leq k$, we obtain two prime factorizations of the number bp; one on the left containing p, one on the right containing primes all smaller than p. As

$$bp = r_1 r_2 \dots r_k < q_1 q_2 \dots q_k = n,$$

we obtained a number smaller than n with two different prime factorizations. This is a contradiction. □

6 What is a proof?

6.1 Formal proofs

A formal mathematical proof is a list of logical formulas (written is a formal language, such as described in Section 3), where each formula is either accepted as true or is a consequence of earlier statements by a rule of inference. The last item on the list is the assertion that has been proved. An axiom is an accepted true statement. An axiom can be an axiom of logic or an axiom specific to a theory; the latter is called a mathematical axiom. For example, the Axiom of Extensionality (Axiom 4.1) is a mathematical axiom.

As for rules of inference, they are different ways of describing them, since all except the first one can be replaced by axioms schemes of logic (an axiom scheme is a list of infinitely many axioms,

following a pattern). We will use three of them. A rule will be described in the form $\frac{\text{premises}}{\text{conclusion}}$. Here, the premises will be a list of one or more formulas, and the conclusion will be a single formula. A rule means that if the premises are established, one can establish the conclusion. In other words, recalling that a proof is a list of formulas, if the premises are already on the list, the rule of inference being used allows one to place the conclusion at the end of the current list (which will be extended until the proof is finished).

Rules of inference

(1) Modus ponens: $\frac{A,\ A \to B}{B}$, where A and B are formulas.

(2) Specialization: $\frac{\forall x A(x)}{A(t)}$, where the term t is substitutable for x in the formula $A(x)$ (see Subsubsection 3.5.1).

(3) Generalization: $\frac{A(x)}{\forall x A(x)}$, where $A(x)$ is a formula and x is a variable that is not free in any of the assumptions.

Rule (3) needs some explanation. As for axioms, mathematical or logical, insofar as they contain free variables, those variables are meant to refer to any object in the theory, that is, they are meant to be universally quantified, but for technical reasons the axioms may be easier to work with if the variables are kept free. On the other hand, assumptions may come with free variables that refer to certain objects satisfying the requirements stated in the assumptionsi (so they are not meant to be universally quantified). For example in Euclid's Lemma 5.1, p, a, and b are free variables in the assumptions. What is meant by Rule (3) is that if we established $A(x)$ without making any assumption about x, then we established $(\forall x)A(x)$. In fact, this rule is involved when one wants to prove something for all integrs n, and one starts out with saying: let n be an arbitrary integer.

As for the axioms of logic, we do not want to say anything; occasionally we may quote such an axiom in a context where such an axiom is natural, but we do not want to make this presentation too formal.

6.1.1 Symbol for provability

Let L be a list of formulas and A a single formula. To state that using the assumptions L there is a proof of A is denoted by the symbol

$$L \vdash A. \tag{6.1}$$

6.2 Computers doing formal proofs

Formal proofs are important both in mathematical logic and in computing, although what computers can do in the way of formal theorem proving is fairly limited, but sometimes very useful, and often exceeds human capabilities. An example of such an impressive achivement is proving a version of the Linux microkernel correct; this is worth a lot of money in the financial industry. Banks do not like their computer systems to hacked, so they need high reliability in their computer programs. In fact, the L4 Linux microkernel (the central part of a version of the Linux operating system) has been proven correct. What has been done was the following. The microkernel L4 has been described in the computer language Haskell, named after the Pennsylvania State Universiy mathemaician

Haskell Curry. This is a human-friendly language, quite unsuitable for writing operating systems. The kernel has been written in C, and what has been proved is that the C implementation does exactly the same thing as the one written in Haskell.

Computers of course do not understand the language C. So, before they can do anything they have to have something to translate the C program into their own machine language. There is a small program, perhaps a thousand lines long, that translate a small part of the C compiler into the computer's own language. This C compiler can translate the rest of the C compiler from C to machine language, and when the compiler has been translated, the rest of the operating system into machine language. Thus the expression "pull oneself up by one's bootstraps" was applied to computers, as in "one bootsraps a computer." Today, one uses the expression to "boot a computer."an expression quite common today.

Thus, to port (implement on another computer) the operating system one only has to rewrite the thousand or so lines on a different computer, rather than rewriting the whole operating system.

6.3 Proofs in mathematical essays

One would very rarely present a formal proof in a mathematical essay. Such proofs are hard to read, and you may need to have a certain inclination to enjoy reading them. A human readable proof at best tries to convice the reader that such a formal proof can be carried out; usually, this objective is not stated clearly, since many mathematicians have only limited training in mathematical logic. Noevertheless, they have a good intuitive understanding when a proof is correct.

6.4 Can one learn how to find proofs?

Fermat's Last Theorem says the following:

Theorem 6.1 (Fermat's Last Theorem)**.** *Given an integer $n > 2$, there are no positive integers a, b, and c such that*

$$a^n + b^n = c^n.$$

The result was stated by Fermat in 1637; he also have claimed to have found a miraculous proof (demonstrationem mirabilem)[6.1] of the result, he did not publish a proof, and a proof was finally found by Andrew Wiles in 1994. One definitely cannot learn how to do such a thing. One can get better at producing proofs, but one will never be perfect at it.

7 Various patterns of mathematical proofs

Often, human presentable mathematical proofs follow certain patters. These patters are not necessarily reflected in the formal version the proof.

7.1 Finding a proof versus presenting it

The final presentation of a proof may be very different from finding it. After finding a proof by a messy and circuitous route, it is customary to clean up the proof and presenting the a simplified version. After this, the proof may be very different from the way it looked originally, and often, motivation is lost. One reads a beautifully written proof, and cannot imagine how anyone ever was able to find such a proof. This is the way one often finds a result, and while one can learn

[6.1] Latin accusative, in a wider context.

mathematics this way, it is difficult to learn how to find proofs by reading such a perfectly polihed diamond.

For this reason, in addition to studying existing mathematical proofs, one may also want to study proof strategies. While many mathematicians develop such strategies on their own, studying strategies that others found useful may speed up the process. In the book "How to solve it" [22], George Pólya identified several problem solving strategies, and initiated *heuristics*,[7.1] the science of problem solving methods. As a take off on the title of the book by Pólya, "How to prove it" by Daniel Velleman [24] is a book discussing ways of finding mathemaical proofs. Both books are excellent supplements to the material in this course.

7.2 Using an intermediate assumption

Given a list of assumptions L, one wants to show A, in symbols $L \vdash A$. In order to do this, one guesses an intermediate assumption B, which then she can use to establish A; formally, $L, B \vdash A$. Then she needs to show B, formally $L \vdash B$. If she succeeds, she established A with the assumptions in L: it was unnecessary to add B to the assumptions, since it can be proved from A. That is, using the proof of B, one can place B into the proof sequence, and attach the proof of A with the aid of B to the end of this proof sequence.

7.3 Direct proof

In a direct proof, one starts out with the assumptions, and step-by-step establishes the result.

7.4 Proof by cases

Given a list of assumptions L, one proves a disjunction $B_1 \vee B_2$. Then one establishes both $L, B_1 \vdash A$ and $L, B_2 \vdash A$. This can be changed to estabslish $L, B_1 \vee B_2 \vdash A$.[7.2] Since we also have $L \vdash B_1 \vee B_2$, the intermediate assumption $B_1 \vee B_2$ can be eliminated, resulting in $L \vdash A$. Instead of the disjunction $B_1 \vee B_2$, we might use a longer disjunction $\bigvee_{i=1}^{n} B_i$, and show $L, B_i \vdash A$ for each B_i.

Often, in proofs of divisibility, the cases B_1 and B_2 are as simple as "n is even" and "n is odd."

7.5 Indirect proof

In an indirect proof, one assumes that the statement to be proved is false, and arrives a contradiction, i.e., a statement that is obvously false. Since one cannot prove a false statement by correct arguments using correct assumptions,[7.3] we must conclude that the assumption that the statement to be proved is false is wrong. Thus, the statement must be true, and the proof is finished.

[7.1] The name is based on the word *Eureka* or *Heureka*, meaning "I found it." the famous explamation attributed to Archimedes, who, after finding his famous hydrostatic law (a body immersed in water loses weight equal to the weight of water displaced), stepped out of the bathtub, and ran naked though the streets of Syracuse, Sicily, eager to share his discovery.

[7.2] This should be clear by common sense. The formal explanation is somewhat complicated, and we will skip it here. The curious reader might be spurred on to study the matter further.

[7.3] At least, one hopes so. The question whether a false statement is a deep mathematical question that has been shown to be unsolvable by Kurt Gödel, though for number theory, there is a proof Gerhard Gentzen that shows one cannot obtain a contradiction. Since the proof uses transfinite induction, it does not contradict Gödel's result.

7.6 Induction: stepping up

Suppose we want to prove a statement $\phi(n)$ in the realm of nonnegative integers $\mathbb{N} = \{0, 1, 2, 3, \ldots\}$.[7.4] Here n is a distinguished variable in $\phi(n)$, but $\phi(n)$ may have other free variables. We proceed to establish $\phi(0)$, and show that $\forall n\big(\phi(n) \to \phi(n+1)\big)$. Having done this, we have established $\forall n\phi(n)$.

The conclusion is natural and needs no justification. In fact, since $\phi(0)$ and $\forall n\big(\phi(n) \to \phi(n+1)\big)$, we can conclude $\phi(1)$ by using the latter for $n = 0$. Then, using it for $n = 1$, we can conclude $\phi(2)$. Continuing this in this way, we can conclude $\phi(3)$, $\phi(4)$, ... By stepping through all the integers, we can conclude that $\forall n\phi(n)$. That this line of argument is correct is expresses by the formula

$$\Big(\phi(0) \,\&\, \forall n\big(\phi(n) \to \phi(n+1)\big)\Big) \to \forall n\phi(n). \tag{7.1}$$

This is added as an axiom for every formula ϕ and one member of the *induction axiom scheme.*[7.5] As we mentioned, $\phi(n)$ may free variables other than n. As we mentioned, an axiom meant to be true for all values of the free variables in it. This is reflected in the way Rule (2) in Subsection 6.1 was stated, where the variable x as allowed to be free in any of the axioms.[7.6]

7.7 Induction: true if true for smaller integers

Sometimes, to establish a formula $\forall n\psi(n)$, one first shows that $\psi(n)$ follows if $(\forall k < n)\psi(k)$. Formally,

$$\forall n\big((\forall k < n)\psi(k) \to \psi(n)\big) \to \forall n\psi(n). \tag{7.2}$$

This, however, does not need to be added as an axiom, since it is a consequence of the induction scheme (7.1) with $\phi(n) = (\forall k < n)\psi(k)$.

Indeed, $\psi(0)$ is *vacuously* (that is, it says nothing, so it is true), as we are about to explain: By untangling the restricted quantifiers (see formula (3.6))

$$\phi(0) \equiv (\forall k < 0)\psi(k) \equiv \forall k\big(k < 0 \to \psi(k),$$

and the right-hand side has a conditional where the antecedent $(k < 0)$ is false, since we are working only with nonnegative integers, and so the conditional is true. The word "vacuously" above meant that for the assertion to be true, no requirement was imposed on $\psi(k)$. Further, we have

$$\phi(n) \to \phi(n+1) \equiv \big((\forall k < n)\psi(k)\big) \to \big((\forall k < n+1)\psi(k)\big)$$

and the right-hand side can be proved from the formulas $k < n+1 \leftrightarrow (k < n \vee k = n)$ and the formula $(\forall k < n)\psi(k) \to \psi(n)$. This shows that formula (7.2) is indeed a consequence of (7.1).

7.7.1 Induction: failure at smallest integers

Often, one combines this form of induction with an indirect proof. That is, one wants to prove that $\forall n\psi(n)$ and assumes that this is not true. So we have $\neg\psi(n)$ for some n. Let n be the smallest integer n for which we have $\neg\psi(n)$. We then show that this is not possible: we do this by showing that $(\forall k < n)\psi(k) \to \psi(n)$. Then $\forall n\psi(n)$ follows by the induction scheme (7.2).

[7.4] The set $\mathbb{N}$ is called the set of natural numbers. In various contexts, 0 sometimes counts as a natural number, at other times it does not. In logic, 0 always considered a natural number.

[7.5] An axiom scheme is a list, usually infinite, of axioms following a certain pattern

[7.6] This rule allows one to restate any axiom with all its free variables universally quantified.

7.8 Second order justification of mathematical induction

The wellordering[7.7] principle for integers says that every nonempty set of nonnegative integers has a least element. In [2, p. 153], this is used to justify induction. We want to avoid such a justification, since the wellording principle is a principle of second order arithmetic (i.e., quantification over all subsets of $\mathbb{N}$ is allowed; see Subsection 3.8 about first order logic). It is a principle much too strong for our purposes, and it involves deep philosiphical questions.

As far as induction is concerned, instead of the full strength of the wellordering principle is concerned, we only need the fact that for each formula $\phi(n)$, the set

$$\{n \in \mathbb{N} : \neg\phi(n)\} \tag{7.3}$$

either has a least element or is empty, and this can be formulated as a first order axiom scheme, basically equivalent to the schemes in (7.1) and (7.2). Thus, the wellordering principle should be restricted to definable sets described in equation (7.3).

This makes a big difference. There are uncountably many subsets of $\mathbb{N}$, as will be discussed later in Section 19 on cardinalities. The cardinality of the set of all subsets of $\mathbb{N}$ leads to currently undecicable problems in axiomatic set theory.

In any case, mathematical induction does not need any justification. It is just as obvious as anything that can be invoked for its justification.

8 Simple examples of proofs

8.1 Trivial situations

Mathematicians often like to use the word trivial to designate a statement that you can immediately ascertain the truth of. The word is not derogatory at all in mathematics, but its use in other areas may be so. It needs to be taken with a grain of salt (i.e., often it cannot be taken literally), since occasionally you might meet a statement the author calls trivial, but you need to give quite a bit of thought to see why this is so.

8.1.1 Vacuous assertions

Sometimes you are faced with an assertion where there is nothing to prove since the assumptions are always false. So, the first task in trying to establish a statement is often to figure out what the assumptions mean.

Problem 8.1. Let n be a positive integer such that $2n + 3/n < 5$. Show that $2n^2 + 5/n^2 < 16$.

Solution. The problem here is that the assumption $2n + 3/n < 5$ is never true. To show this is a simple exercise in calculus. The derivative of the function $f(x) = 2x + 3/x$ is $f'(x) = 2 - 3/x^2$, and this is positive if $x \geq 1$; so $f(x)$ is increasing if $x \geq 1$. Now $f(1) = 5$, so $f(n) < 5$ is false for all positive integers. We are asked to prove

$$(\forall n \geq 1)\left(2 + \frac{3}{n} < 5 \to 2n^2 + \frac{5}{n^2}\right),$$

but the antecedent $2n + 3/n < 5$ of the conditional is false for all $n \geq 1$, so the statement is true. One would say that the statement is *vacously* true (true in an empty way), since it really asserts nothing.

[7.7] Also written as well-ordering. A set is said to be wellordered (by some order relation) if every subset has a first element.

8.1.2 Unnecessary assumptions

Sometimes, an assertion is complicated by unnecessary assumptions, since the assertion is true without the assumptions.

Problem 8.2. Assume n is an odd integer. Show that $2n^2 + 4$ is even.

Solution. Here the assumption that n is odd is unnecessary, so we will ignore this assumption. We have $2n^2 + 4 = 2(n^2 + 2)$, which is twice an integer, so it is even.

8.2 Simple direct proofs

Problem 8.3. Let n be an integer. Show that $n^2 + n$ is even.

It is not assumed here that n is positive; we may have $n = 0$ or $n < 0$.

Solution. We have $n^2 + n = n(n + 1)$. One of n and $n + 1$ is even, so their product is even.

Sometimes, when dealing with a number of problems, it helps to remembr what we did before.

Problem 8.4. Let n be an integer. Show that $n^2 + n + 1$ is odd.

Solution. We just showed that $n^2 + n$ is even. So $(n^2 + n) + 1$ must be odd.

Problem 8.5. Prove that the product of two consecutive positive integers is not a square of an integer.

Solution. Given a positive integer n, its square is n^2, and the square of the next integer is $(n+1)^2 = n^2 + 2n + 1$. Since $n(n + 1) = n^2 + n$ is between these two numbers, it cannot be the square of an integer.[8.1]

Problem 8.6. Prove that an integer n is the sum of the squares of two integers if and only if $2n$ has the same property (i.e., $2n$ is also the sum of the squares of two integers; note that one or both of these integers may or may not be zero).

Solution. If $n = x^2 + y^2$ for some integers x and y then $2n = 2x^2 + 2y^2 = (x + y)^2 + (x - y)^2$. This shows that if n can be represented as the sum of squares of two integers then $2n$ can also be so represented.

If, on the other hand, $2n = u^2 + v^2$ for some integers u and v then $n = x^2 + y^2$ with x and y such that $u = x + y$ and $v = x - y$, that is, with

$$x = \frac{u + v}{2} \quad \text{and} \quad y = \frac{u - v}{2}.$$

Observing the equation $2n = u^2 + v^2$ implies that u and v have the same parity (i.e., that either they are both even or they are both odd), it follows that x and y are integers. Hence it also follows that if $2n$ can be represented as the sum squares of two integers then n can also be so represented.[8.2]

Problem 8.7. Let $n \geq 2$ be an integer, and let $a_1, a_2, \ldots, a_n$ be nonzero real numbers, and assume $a_n = a_1$. Show that there is an even number of integers k with $1 \leq k \leq n - 1$ for which $a_k a_{k+1} < 0$.

[8.1] See [13, Problem 1] for the source of the problem.
[8.2] See [13, Problem 3] for the source of the problem.

Solution. Noting that $a_n = a_1$, we have

$$\prod_{k=1}^{n-1} a_k a_{k+1} = \prod_{k=1}^{n-1} a_k^2 > 0;$$

therefore, there must be an even number of negative factors in the product on the left-hand side.[8.3]

8.2.1 Brute force versus beauty

Problem 8.8. Show that the product of four consecutive integers plus 1 is always a square of an integer.[8.4]

Solution. The main challenge in the problem is to obtain the result without messy calculations. Let the four consecutive integers be $k-1$, k, $k+1$, and $k+2$. Then we have

$$\begin{aligned}(k-1)k(k+1)(k+2) &= (k-1)(k+2)\,(k^2+k) = (k^2+k-2)(k^2+k)\\ &= \big((k^2+k-1)-1\big)\big((k^2+k-1)+1\big) = (k^2+k-1)^2-1,\end{aligned}$$

showing that indeed

$$(k-1)k(k+1)(k+2)+1 = (k^2+k-1)^2$$

is the square of an integer.

Choosing the notation here simplified the calculation. If we had denoted the four integers as k, $k+1$, $k+2$, $k+3$, the solution would have been more difficult to find.

8.3 Proof by contrapositive

We have

$$(A \to B) \leftrightarrow (\neg B \to \neg A),$$

The conditional on the right is called the contrapositive of this on the left. We can make use of this tautology in some proofs.

Problem 8.9. Let n be an integer and assume $n^2 - 2n + 3$ is even. Show that n is odd.

Solution. We need to prove the conditional "if $n^2 - 2n + 3$ is even then n is odd." Instead, we prove the contrapositive "if n is not odd then $n^2 - 2n + 3$ is not even," which says the same thing in a different way.

Stated more simply, we need to prove that if n is even then $n^2 - 2n + 3$ is odd. Assuming n is even, let k be an integer that $n = 2k$. Then

$$n^2 + 2n + 3 = (2k)^2 - 2(2k) + 3 = 2(2k^2 - 2k + 1) + 1,$$

and the righ-hand side is clearly odd.

[8.3] See [12, Problem 2] for the source of the problem.
[8.4] See [19, Problem 1] for the source of the problem.

8.4 Proof by cases

Problem 8.10. Let n be an integer. Show that $n^2 - 3n + 7$ is odd.

Solution. While there are different ways of attacking this problem, one way to do this is to distinguish to cases: 1) n is even, or 2) n is odd, and prove the assertion in each of these case.

Case 1. Assume n is even. Let k be an integr such that $n = 2k$. Then

$$n^2 - 3n + 7 = (2k)^2 - 3(2k) + 7 = 4k^2 - 3 \cdot 2k + 6 + 1 = 2(2k^2 - 3k + 3) + 1,$$

and the right-hand side is odd,

Case 2. Assume n is odd. Let k be an integer such that $n = 2k + 1$. Then

$$\begin{aligned} n^2 - 3n + 7 &= (2k + 1)^2 - 3(2k + 1) + 7 = (4k^2 + 4k + 1) - 6k - 3 + 7 \\ &= 2(2k^2 - k + 2) + 1, \end{aligned}$$

and the right-hand side is again odd.

8.5 Problems involving real numbers

The next examples show that if the problem has certain symmetries, it is often important to preserve these symmetries in the solution; this is, however, not always possible.

Problem 8.11. Given real numbers a, b, and c, show that

$$a^2 + b^2 + c^2 \geq ab + bc + ca.$$

Solution. We have

$$0 \leq (a - b)^2 + (b - c)^2 + (c - a)^2 = 2(a^2 + b^2 + c^2 - ab - bc - ca),$$

whence the assertion follows. Clearly, equality holds only in case $a = b = c$.[8.5]

Problem 8.12. Let $A \neq 0$, $B \neq 0$, $C \neq 0$ and a, b, c be real numbers, and assume that

$$a + b + c = 0, \qquad A + B + C = 0, \qquad \frac{a}{A} + \frac{b}{B} + \frac{c}{C} = 0.$$

Show that

$$aA^2 + bB^2 + cC^2 = 0.$$

Solution. Writing

$$S = aA^2 + bB^2 + cC^2,$$

and noting that we have $A = -B - C$, $B = -A - C$, and $C = -A - B$ according to the second of the given equations, we obtain that

$$\begin{aligned} 2S = S + S &= \big(aA^2 + bB^2 + cC^2\big) + \big(a(B + C)^2 + b(A + C)^2 + c(A + B)^2\big) \\ &= a(A^2 + B^2 + C^2) + b(A^2 + B^2 + C^2) + c(A^2 + B^2 + C^2) + 2aBC + 2bAC + 2cAB \\ &= (a + b + c)(A^2 + B^2 + C^2) + 2(aBC + bAC + cAB). \end{aligned}$$

The first term on the right-hand side is 0 in view of first of the given equations, and the second term is 0 as can be seen by multiplying the third of the given equations by ABC. Hence $S = 0$, as we wanted to show.[8.6]

[8.5] See [15, Problem 2] for the source of the problem.
[8.6] See [21, Problem 1] for the source of the problem.

8.6 Equality of sets

Problem 8.13. Let A, B, C be sets. Show that

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C).$$

Solution. Let x be an arbitrary element of at least one of these sets. Defining the logic formulas $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{C}$ as $\mathcal{A} = x \in A$, $\mathcal{B} = x \in B$, and $\mathcal{C} = x \in C$, we have

$$x \in A \cap (B \cup C) \leftrightarrow \big(\mathcal{A} \,\&\, (\mathcal{B} \vee \mathcal{C})\big)$$

and

$$x \in (A \cap B) \cup (A \cap C) \leftrightarrow \big((\mathcal{A} \,\&\, \mathcal{B}) \vee (\mathcal{A} \,\&\, \mathcal{C})\big)$$

according to the definition of the set operations involved. In view of the Axiom of Extensionality (Axiom 4.1), we need to show that x is an element of the first set if and only if it is an element of the second. This immediately follows from the observation that the logic formula

$$\big(\mathcal{A} \,\&\, (\mathcal{B} \vee \mathcal{C})\big) \leftrightarrow \big((\mathcal{A} \,\&\, \mathcal{B}) \vee (\mathcal{A} \,\&\, \mathcal{C})\big)$$

is a tautology, as can easily be verified by its truth table.

The solution of this problem presents a general method of transforming certain tautologies involving a biconitional into equalities involving set operations.

8.7 Reading

[2, Chapter 3, pp. 81–104], [2, §4.3–4.6, pp. 113–126].

There is lot here to read and do homework; you can stretch this out a little while we are continuing with the notes. The same applies to the homework.

8.8 Homework

[2, Chapter 3, pp. 81–104], p. 84, 3.1, 3.5, 3.7, p. 88, 3.9, 3.11, 3.15, p. 94, 3.17, 3.21, p. 97, 3,27, 3.29, 3.37. p. 102, 3,49, 3.59. [2, §4.3–4.6, pp. 113–126], p. 116, 4.29, 4.33, 4.37, p. 120, 4.41, 4.45. 4.49, p.122, 4.53, 4.55, 4.49, p. 123, 4.61, 4.63, 4.67/

9 Counter examples and indirect proofs

9.1 Counterexamples

One can disprove an assertion of form $\forall x A(x)$ by exibiting a term t for which $\neg A(t)$, assuming t is substitutable for x in $A(x)$ (see Subsubsection 3.5.1), i.e., an example for which the assertion $A(x)$ fails. Such an example is called a *counterexample*. A similar situtation can occur with more than one universally quantified variables.

9.1.1 An algebra mistake

Students of introductory algebra occasionally make the mistake and write $\sqrt{a+b} = \sqrt{a} + \sqrt{b}$. It is easy give a counterexample to show that this is not always true (in fact, it is almost never true). For example,

$$5 = \sqrt{25} = \sqrt{9+16} \neq \sqrt{9} + \sqrt{16} = 3 + 4 = 7.$$

9.1.2 Euler's power sum conjecture

In analogy with Fermat's Last Theorem (see Theorem 6.1), in 1769 Euler conjectured that, given $n > 2$, to obtain the nth power of an integer, one needs to sum at least n powers of positive integers. In other words, the equation

$$\sum_{i=1}^{k} a_k^n = b^n$$

cannot hold for k with $2 \leq k < n$ and positive integers b and a_i $(1 \leq i \leq k)$. The conjecture was disproved by L. J. Lander and T. R. Parkin [6]. Using computers, they found that

$$27^5 + 84^5 + 110^5 + 133^5 = 144^5.$$

Later, counterexamples were found for other exponents, including $n = 4$. For $n = 3$, the assertion is true; in fact it is the case $n = 3$ of Fermat's Last Theorem (the case $n = 3$ of Fermat's Last Theorem was essentially proved by Euler in 1770 (the proof had some gaps, but the gaps could be filled by Euler's own work).

9.2 Indirect proofs

In indirect proofs, one assumes that the result to be proved is false, and using this assumption, one produces a contradiction. Using correct arguments, the only reason to obtain a contradiction is having made a false assumption. Thus the result is established.

9.2.1 A number that is never a square

Problem 9.1. Prove that $n^4 + 3n^2 + 2$ is never a square of an integer.

Solution. Assume, on the contrary, that k is a positive integer such that the given number is equal to k^2. We then have

$$k^2 = n^4 + 3n^2 + 2 = (n^2 + 1)(n^2 + 2),$$

and so k must be between $n^2 + 1$ and $n^2 + 2$, showing that k cannot be an integer.[9.1]

9.2.2 A nonexistent triangle

Problem 9.2. Show that there is no triangle whose altitudes are of length 4, 7, and 10 units.

Solution. Assume there is such a triangle. Let the sides of the triangle be a, b, and c, and let the corresponding altitudes be 4, 7, and 10. Writing A for the area of the triangle, we then have

$$2A = 4a = 7b = 10c,$$

that is

$$b = \frac{4}{7}a \quad \text{and} \quad c = \frac{4}{10}a = \frac{2}{5}a.$$

Then

$$b + c = \left(\frac{4}{7} + \frac{2}{5}\right) a = \frac{34}{35}a < a.$$

This is a contradiction, since in any triangle, the sum of two sides are greater then the third side.[9.2]

[9.1] See [8, Problem 1] for the source of the problem.
[9.2] See [10, Problem 3] for the source of the problem.

9.2.3 Numbers not in a geometric progression

Problem 9.3. Show that $\sqrt{2}$, $\sqrt{5}$, and $\sqrt{7}$ cannot belong to the same geometric progression.

Solution. Assuming that these numbers belong to the same geometric progression $a_n = aq^n$, $n = 0$, 1, 2, ..., where $a > 0$, $q > 0$, and $q \neq 1$, we have $\sqrt{2} = aq^{k_1}$, $\sqrt{5} = aq^{k_2}$, $\sqrt{7} = aq^{k_3}$ for some nonnegative integers k_1, k_2, and k_3. Dividing the second equation by the first equation, we obtain $\sqrt{5/2} = q^{k_2-k_1}$, i.e., $5/2 = q^{2(k_2-k_1)}$, Similarly, dividing the third equation by the first equation, we obtain $\sqrt{7/2} = q^{k_3-k_1}$, i.e., $7/2 = q^{2(k_3-k_1)}$, Writing $l = k_2 - k_1$, and expressing q from the equation giving $5/2$, we obtain $q = (5/2)^{1/(2l)}$ Then writing $k = k_3 - k_1$ substiting it into the equation expressing $7/2$, we obtain

$$\frac{7}{2} = \left(\left(\frac{5}{2}\right)^{1/(2l)}\right)^{2k} = \left(\frac{5}{2}\right)^{k/l}$$

with $k = k_3 - k_1$ and $l = k_2 - k_1$; note that, clearly, $k_1 < k_2 < k_3$, and so $k > l > 0$, in case $q > 1$, and $k_1 > k_2 > k_3$, and so $k < l < 0$, in case $0 < q < 1$. and $l > 0$. Thus, we have

$$7^l \cdot 2^{k-l} = 5^k.$$

This is impossible in case $k > l > 0$, because no positive integer power of 5 is divisible by 2. In case $k < l < 0$, write this equation equivalently as

$$7^{-l} \cdot 2^{-(k-l)} = 5^{-k},$$

so that all the exponents are positive. This equation is impossible for the same reason.[9.3]

9.2.4 The Sylvester–Gallai theorem

This was a famous problem proposed by Sylvester in 1893, and solved, among others, by Gallai in 1944.

Theorem 9.1. *Given a finite number of points in the plane, assume that every line that goes through at least two points, also goes through a third one. Then all the points lie on a line.*

The following strikingly simple proof of this theorem was given by Leroy Milton Kelly.

Proof. Assume there is such a set of points that do not lie on a single line. Let P be one among the points, and l one among the lines such that the distance from P to l is the smallest possible among all the nonzero distances between the points and the lines. Let Q be the point on l that is closest to P; that is, PQ is perpendicular to l, and PQ is the smallest distance we mentioned. The line l contains at least three points of the given points, so at least two of these points lie on one side of the point Q. Say, these two points are R and S, and assume R lies between Q and S. Then the distance between the points R and the line PS is smaller than PQ. This is a contradiction, since PQ was supposed to be the smallest distance. □

See the Wikipedia article for an extensive discussion of the Sylvester–Gallai theorem, including a figure illustrating the above proof (our notation is different from that used in the figure).

This proof is interesting in that it produces a counterexample to an asseertion by focusing on a least quantity (smallest distance). The proof is, however, not an induction proof, certainly, not induction on the number of points. This illustrates the difficulty of classifying proofs. For example,

[9.3] See [13, Problem 2] for the source of the problem.

the proofs of Euclid's Lemma 5.1 and of the Fundamental Theorem 5.1 are clearly induction proofs, since they focus on the least counterexample failing to satisfy the assertion. On the other hand, the proofs of Lemma 5.2 and Corollaries 5.1 and 5.2 are not induction proofs, even though they focus on a least quantity satisfying a certain requirement.

Yet, all these proofs have strong common features, perhaps a stronger commonality than proofs by inductions have. So, classifying a proof by the least counterexample as an induction proof is apparently not a natural, and probably not a helpful, classification. Virtually the only use of this classification is to show that such a proof, when applied to integers, is sufficiently justified by the induction scheme (7.1) (note that (7.2) is a consequence of the former).

9.3 Reading

[2, Chapter 5, pp. 127–151]. The irrationality of square roots is discussed later, in Sections 14, 15, and 16.

9.4 Homework

[2, Chapter 5, pp. 127–151]. p. 131, 5.3, 5.5, 5.9, p. 137, 5,13, 5.19, 5.27, 5.29, p. 145, 5.43, 5.45, 5.49, p. 149, 5.55, 5.57. p. 33, p. 34, p. 36,

10 Mathematical induction

Induction was discussed in Subsections 7.6 and 7.7, but examples for the use of induction occurred before.

10.1 Recursive definition

In a recursive definition, one defines an instance of an object in term of instances of the same object that have already been defined. For example, for a function for nonnegative integers, one defines $f(0)$, and for $n > 0$, one defines $f(n)$ in terms of $f(n-1)$. For example, the sum

$$f(n) = \sum_{k=1}^{n} k^2$$

can be defined recursively as $f(0) = 0$ and

$$f(n) = f(n-1) + n^2. \tag{10.1}$$

Such recursive definitions can often be used in proof by induction.

Problem 10.1. Show that

$$\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}.$$

for all integers $n \geq 1$.

Solution. Writing $f(n)$ for the sum, the assertion for $n = 0$ is obviously true.[10.1] since $f(0) = 0$. Let $n > 0$, and assume the result is true for $n - 1$ replacing n; that is, assume

$$f(n-1) = \frac{(n-1)\big((n-1)+1\big)\big(2(n-1)+1\big)}{6}. = \frac{(n-1)n(2n-1)}{6}.$$

Using equation (10.1), we then have

$$f(n) = f(n-1) + n^2 = \frac{(n-1)n(2n-1)}{6} + n^2 = \frac{n(n^2-3n+1)}{6} + n^2$$
$$= \frac{n(n^2-3n+1)+6n^2}{6} + n^2 = \frac{n(n^2+3n+1)}{6} = \frac{n(n+1)(2n+1)}{6},$$

showing that the result is also true for n. This establishes the result by induction.

Stepping from $n-1$ to n rather than stepping from n to $n+1$ simplifies the algebraic manipulations needed for establishin the result. Similar sums are extensively discussed in the note [20].

Problem 10.2. Define a sequence as follows: $a_1 = 2$, and $a_{n+1} = a_n^2 - a_n + 1$ for all $n \geq 1$. Show that if $m > n \geq 1$ then a_m and a_n are relatively prime (i.e., their greatest common divisor is 1).

Solution. Let $k > 1$, $n \geq 1$ and assume that $k \mid a_n$ (in words: k is a divisor of a_n) or that $k \mid a_n - 1$. We claim that then $k \mid a_m - 1$ for all $m > n$.

Indeed, this is true for $m = n + 1$, since $a_{n+1} - 1 = a_n(a_n - 1)$. Assume that for some $m > n$ we have $k \mid a_m - 1$. Then $k \mid a_m(a_m - 1) = a_{m+1} - 1$. Hence, our claim follows for all $m > n$ by induction.

Thus if $k > 1$ and $1 < n < m$ and $k \mid a_n$, then $k \nmid a_m$ since $k \mid a_m - 1$.[10.2]

10.2 Reading

[2, Chapter 6, pp. 152–180], [2, Chapter 7, pp. 181–193]. [2, Chapter 8, pp. 200–220].

10.3 Homework

[2, Chapter 6, pp. 152–180]. p. 161, 6.5, 6.9, 6.11, 6.15, p. 169, 6.21, 6.27, 6.31, p. 174, 6.35, 6.37, p. 177, 6.47, [2, Chapter 7, p. 193]. 7.3 (part (a) of this problem is discussed above as Problem 8.8 above) 7.11 (Pythagerean triples are discussed below in Subsections 12.1 and 12.2), [2, Chapter 8, pp. 200–220], p. 209, 8.11, 8.17, 8.23.

11 Equivalence relations

11.1 Cartesian products

Given a set the sets A and B, their Cartesian product is the set

$$A \times B = \{(a, b) : a \in A \text{ and } b \in B\};$$

here (a, b) is an ordered pair whose first member is a and its second membeer is b.[11.1]

[10.1] The assertion is only made for $n \geq 1$, but since it is also true for $n = 0$, we show the result for all $n \geq 0$. It is simpler to show that the result is true for $n = 0$ than to show it for $n = 1$; this is why we start at $n = 0$.

[10.2] See [19, Problem 2] for the source of the problem.

[11.1] An ordered pair of items lumps two items together in a way that on lumps together two items in a way that one can tell which is a first member and which is the second member. The ordered pair of two things a and b is

11.2 Relations

Definition 11.1. Let A and B be sets. A relation R on the sets A and B is a set $R \subset A \times B$. A relation R on the set A is a set $R \subset A \times A$.

For a relation R, instead of saying $(a, b) \in R$ one often says aRb. Given R, one can talk about the domain and the range of a relation.

$$\mathrm{dom}(R) \stackrel{def}{=} \{x : \exists y \, (x, y) \in R\}, \tag{11.1}$$

and

$$\mathrm{ra}(R) \stackrel{def}{=} \{y : \exists x \, (x, y) \in R\}. \tag{11.2}$$

The inverse of a relation results by reversing the pairs:

$$R^{-1} \stackrel{def}{=} \{(x, y) : (y, x) \in R\}. \tag{11.3}$$

Definition 11.2. Let R be a relation on the set A

(1) R is called *reflexive* if xRx for all $x \in A$.

(2) R is called *symmetric* if, given $x, y \in A$, we have xRy if and only if yRx.

(3) R is called *transitive* if, given $x, y, z \in A$ such that xRy and yRz then xRz.

(4) R is called *irreflexive* or *anti-reflexive* if $\neg \, xRx$ for all $x \in A$.

(5) R is called *connected* if, given $x, y \in A$, we have $x = y$ or xRy or yRx.

Instead of saying R is reflexive, one may say R is reflexive on A; similarly for connected. The other properties do not depend on A.[11.2]

11.3 Equivalence relations

Definition 11.3. A relation on a set A that is reflexive, symmetric, and transitive is called an *equivalence relation*. If R is an equivalence relation on A, and $x \in A$, one writes

$$[x]_R \stackrel{def}{=} \{y \in A : xRy\}.$$

As we will show, equivalence relations are associated with partitions.

Definition 11.4. Let A be a set. A set P is called a partition of A if $\bigcup P = A$ and the elements of P are pairwise disjoint sets.

We have

usually denoted as (a, b). In mathematics, there are various representations of ordered pairs; the most widely accepted set-theoretical definition is the Kuratowski representation: $(a, b) = \{\{a\}, \{a, b\}\}$. We will not need to get involved in the subtleties of this representation.

[11.2] If R is a relation on the set A and $A \subset B$, then A is also a relation on B. If $A \neq B$ in addition, then R is not reflexive and not connected on B even if R is reflexive and connected on A. The other properties and not affected by considering R on B.

Theorem 11.1. *If R is an equivalence relation on A, then*

$$P_R \stackrel{def}{=} \{[x]_R : x \in A\}$$

is a partition of A.

Proof. We have $\bigcup P_R \subset A$. Indeed, if $X \in P_R$ then $X = [x]_R$ for some $x \in A$. If $y \in [x]_R$ then xRy, so $y \in A$. We also have $A \subset \bigcup P_R$. Indeed, if $x \in A$, then we have xRx by reflexivity, and so $x \in [x]_R \in P_R$. Therefore $\bigcup P_R = A$.

Furthermore, if $X, Y \in P_R$ and $X \neq Y$ then $X \cap Y = \emptyset$. Indeed, assume, on the contrary, that we have $X \cap Y \neq \emptyset$, and let z be such that $z \in X \cap Y$. We have $X = [x]_R$ for some $x \in A$. Thus, $z \in [x]_R = X$, and so xRz. Similarly, $Y = [y]_R$, for some $y \in A$. Therefore, $z \in [y]_R = Y$, and so yRz. By symmetry, we have zRy. Given that we have xRz and zRy, we have xRy by transitivity. We also have yRx by symmetry.

It is now easy to show that $X \subset Y$. Indeed, let $u \in X$ be arbitrary. Then $u \in [x]_R = X$, and so xRu. Since we showed that yRx, this implies yRu by transitivity. Thus, $u \in [y]_R = Y$. This shows that $X \subset Y$. A similar argument, with the roles of X and Y interchanged, shows that $Y \subset X$. Hence $X = Y$. □

If R is an equivalence relation, the sets $[x]_R$ are called *equivalence classes* with respect to R. The above proof shows that if $y \in [x]_R$ then $[x]_R = [y]_R$. If $y \in [x]_R$, then y is called a representative of the equivalence class $[x]_R$. Conversely, we have the following.

Theorem 11.2. *Let A be a set and let P a partition of A. Then the relation R_P defined as*

$$xR_Py \leftrightarrow (\exists X \in P)(x, y \in X)$$

is an equivalence relation.

The proof is left to the reader.

11.4 Reading

[2, §9.1–9.4, pp. 224–238].

11.5 Homework

[2, §9.1–9.4, pp. 224–238], p. 225, 9.1, 9.5, p. 229, 9.11, 9.17, 9.23, p. 234, 9.25, 9.29, 9.33, p. 239, 9.37, 9.39, 9.43.

12 Divisiblity

Definition 12.1. Given two integers d and n, we say that d is a *divisor* of n if there is an integer k such that $n = kd$.

In symbols, one says $d \mid n$ to express that d is a divisor of n. One expresses the same also by saying that d is a *factor* of n, or n is *divisible* by d, or n is a multiple of d. The symbol $d \nmid n$ means that it is not the case that $d \mid n$. Note that $0 \mid n$ is true only if $n = 0$, and $n \mid 0$ is true for all integers n. We have

Lemma 12.1. *Let a, b, c be integers. 1) If $a \mid b$ then $a \mid bc$. 2) If $a \mid b$ and $b \mid c$ then $a \mid c$. 3) If $a \mid b$ and $a \mid c$ then for all integers x and y we have $a \mid bx + cy$.*[12.1]

The result is given in [2, pp. 105–106], except that in the definition of $a \mid b$ the assumption $a \neq 0$ is made, and this assumption necessitates a more complicated statement of the result. The proof is immediate from the definition.

Proof. 1) By the assumption we have $b = ak$ for some integer k, and so $bc = a(ck)$.

2) By the assumptions, there are integers k and l, such that $b = ak$ and $c = bl$, and so $c = a(kl)$.

3) By the assumptions, there are integers k and l such that $b = ka$ and $c = la$, and so $bx + cy = a(kx + ly)$. □

In the next lemma, (k, l) denotes the greatest common divisor of the integers k and l. The greatest common divisor is discussed in Section 17.

Lemma 12.2. *Let k, l and n be integers such that $(k, l) = 1$. If $k \mid n$ and $l \mid n$ then $kl \mid n$.*

Proof. We have $k \mid n$, so $n = ka$ for some integer a. We have $l \mid n$, so $l \mid ka$. Since $(k, l) = 1$, we have $l \mid a$ by Corollary 5.2; i.e, $a = lb$ for some integer b. Hence $n = b(kl)$, and so $kl \mid n$. □

Problem 12.1. Prove that if n is an even integer then $n(n + 1)(n + 2)$ is divisible by 24.[12.2]

Solution. One of n, $n + 1$, and $n + 2$ is divisible by 3, so $n(n + 1)(n + 2)$ is divisible by 3. We need to show that it is also divisible by 8. Let $n = 2k$, where k is an integer. Then k or $k + 1$ is even, so $k(k + 1)$ is divisible by 2. Hence $n(n + 2) = 4k(k + 1)$ is divisible by 8, as we wanted to show.

Problem 12.2. Show that for any integer n, the number $n^3 + 11n$ is divisible by 6.

Solution. We need to show that $n^3 + 11n$ is divisible both by 2 and 3. We have

$$n^3 + 11n = n(n^2 + 11).$$

This is clearly divisible by 2, since for even n the first factor is even, and for odd n the second factor is so.

As for divisibility by 3, it is enough to show that

$$(n^3 + 11n) - 3 \cdot 4n = n^3 - n = (n^2 - 1)n = (n - 1)n(n + 1)$$

is divisible by 3. This is certainly true, since one of the factors on the right-hand side is divisible by 3. [12.3]

Problem 12.3. Given a positive integer n, show that $10n^3 + 3n^2 - n$ is divisible by 6.

Solution. We have

$$10n^3 + 3n^2 - n = 2(5n^3 + n^2) + n(n - 1);$$

since either n or $n - 1$ is divisible by 2, the left-hand side also must be divisible by 2. Further, we have

$$10n^3 + 3n^2 - n = 3(3n^3 + n^2) + n(n - 1)(n + 1);$$

since one of n, $n - 1$, and $n + 1$ is divisible by 3, the left-hand side must be divisible by 3. Thus $10n^3 + 3n^2 - n$ is divisible by both 2 and 3, and so it is divisible by 6.[12.4]

[12.1] One might be tempted to use parentheses here and write $a \mid (bx + cy)$, in analogy with $(bx + cy)/a$, but this is unnecessary, since $(a \mid bx) + cy$ makes no sense, so there is no possibility of misunderstanding.

[12.2] See [7, Problem 1] for the source of the problem.

[12.3] See [18, Problem 1] for the source of the problem.

[12.4] See [9, Problem 1] for the source of the problem.

Problem 12.4. Let n be an integer. Prove that

$$n^4 + 6n^3 - n^2 + 18n$$

is divisible by 24.

Solution. Writing

$$N = n^4 + 6n^3 - n^2 + 18n,$$

we need to show that N is divisible by 3 and by 8. We have

$$N = (n^4 - n^2) + 3(2n^3 + 6n) = n \cdot (n-1)n(n+1) + 3(2n^3 + 6n).$$

Since one of the numbers $n-1$, n, and $n+1$ is divisible by 3, their product is also divisible by 3. Hence N is divisible by 3.

Furthermore,

$$\begin{aligned} N &= (n^4 - 2n^3 - n^2 + 2n) + 8(n^3 + 2n) = (n^3(n-2) - n(n-2)) + 8(n^3 + 2n) \\ &= (n^3 - n)(n-2) + 8(n^3 + 2n) = (n-2)(n-1)n(n+1) + 8(n^3 + 2n). \end{aligned}$$

There are two even numbers among the numbers $n-2$, $n-1$, n, $n+1$, and one of these is divisible by 4. Thus, their product is divisible by 8, showing that N is also divisible by 8, which is what we wanted to show.[12.5]

Problem 12.5. Show that the sum of the cubes of three consecutive integers (i.e., integers that are adjacent, or following one another) is divisible by 18 if and only if the middle integer is even.

Solution. Let the integers be $n-1$, n, and $n+1$.[12.6] Then

$$(n-1)^3 + n^3 + (n+1)^3 = 3n^3 + 6n = 3n(n^2 + 2).$$

In order for this to be divisible by 18, we need to make sure that $n(n^2+2)$ is divisible by 6. Since n and n^2+2 have the same parity, $n(n^2+2)$ is divisible by 2 if and only if n is even.

On the other hand, $n(n^2+2)$ is always divisible by 3. Indeed, this is certainly the case if n is divisible by 3. Assume this is not the case. Then $n = 3k \pm 1$ for some k, and so

$$n^2 + 2 = (9k^2 \pm 6k + 1) + 2 = 9k^2 \pm 6k + 3,$$

which is divisible by 3. So

$$(n-1)^3 + n^3 + (n+1)^3$$

is divisible by 18 if and only if n is even. In other words, the sum of the cubes of three consecutive integers is divisible by 18 if and only if the middle number is even.[12.7]

Note: To show that n^2+2 is divisible by 3 in case n is not divisible by 3, one can also argue that $n^2 + 2 = (n^2 - 1) + 3$ is divisible by 3 according to Fermat's Theorem 13.1.

Problem 12.6. Prove that the sum of the squares of five consecutive integers is divisible by 5, but it is not divisible by 25.

[12.5] See [12, Problem 1] for the source of the problem.
[12.6] See Subsubsection 8.2.1 why we did not chose the integers to be n, $n+1$, and $n+2$.
[12.7] See [17, Problem 3] for the source of the problem.

Solution. Let the five integers be $n-2$, $n-1$, n, $n+1$, and $n+2$. Then

$$(n-2)^2+(n-1)^2+n^2+(n+1)^2+(n+2)^2 = 5n^2+10 = 5(n^2+2),$$

and this is clearly divisible by 5. To show that it is not divisible by 25, we need to show that n^2+2 is not divisible by 5. For this, we need to examine the cases $n = 5k$, $n = 5k \pm 1$, and $n = 5k \pm 2$ for some integer k. In these cases, n^2+2 in turn equals $25k^2+2$, $25k^2 \pm 10k + 3$, $25k^2 \pm 20k + 6$. It is clear that none of these numbers are divisible by 5.[12.8]

12.1 Pythagorean triples

The positive integers a, b, and c such that $a^2+b^2 = c^2$ are said to form a *Pythagorean triple*. We are interested in the form of Pythagorean triples. First note that if a, b, and c form a Pythagorean triple and k is a positive integer, then ka, kb, and kc is also a Pythgorean triple. For this reason, the Pythagorean triples of primary interest will be those for which the greatest common divisor of a, b, and c is 1, i.e., for which $(a,b) = 1$. Such Pythagorean triples are called *primitive*. It is obvious that in a such a triple, a or b must be odd.

Lemma 12.3. *Let the positive integers a, b, c form a primitive Pythagorean triple; that is, $a^2+b^2 = c^2$ and $(a,b) = 1$. Assume that a is odd. Then there are relatively prime positive integers x and y such that $a = x^2-y^2$, $b = 2xy$, and $c = x^2+y^2$.*

Proof. First, note that b must be even. Indeed, if both a and b are odd, say $a = 2k+1$ and $b = 2l+1$, then

$$c^2 = a^2+b^2 = (2k+1)^2+(2l+1)^2 = 4(k^2+k+l^2+l)+2,$$

which is impossible, because the right-hand side is even, but it is not divisible by 4. So $2 \mid c^2$; hence $2 \mid c$ (by Euclid's Lemma 5.1 with $p = 2$, since $2 \mid c \cdot c$), and so $4 \mid c^2$.

Thus, a is odd, b is even, and c is odd. Therefore

$$b^2 = c^2-a^2 = 4\frac{c-a}{2}\frac{c+a}{2}. \tag{12.1}$$

Since both a and c are odd, the fractions on the right-hand side are integers; write $s = (c-a)/2$ and $t = (c+a)/2$.

Next, note that we must have $(s,t) = 1$; indeed, if p is a prime divisior of both s and t, then $p \mid s+t = c.$ and $p \mid t-s = a$, so $p^2 \mid c^2-a^2 = b^2$. Hence $p \mid b$, which contradicts $(a,b) = 1$.

The equation above says that $b^2 = 4st$. Therefore, each of s and t must be a square of an integer. This is because in the prime factorization of b^2 every prime occurs with an even exponenet; since s and t have no common prime factors, each prime in the prime factorization of s or t must occur with an even exponent. So write $s = y^2$ and $t = x^2$ with relatively prime positive integers x and y. Then $b^2 = 4x^2y^2$ according to equation (12.1) and the equations expressing s and t following it. Hence $b = 2xy$, and the latter equations also show that $a = t-s = x^2-y^2$ and $c = t+s = x^2+y^2$. □

12.2 The simplest case of Fermat's Last Theorem

The result describing Pythagorean triples we just showed can be used to establish the simplest case of Fermat's Last Theorem (Theorem proof: Fermat's last, thm):

[12.8] See [21, Problem 1] for the source of the problem.

Theorem 12.1. *There are no positive integers a, b, and c such that*

$$a^4 + b^4 = c^2. \tag{12.2}$$

The proof uses a form of induction described in Subsection 7.7.1. Fermat called the method of proof *descente infinite.*[12.9] This is the only case of Fermat's Last Theorem in which Fermat published a proof.

Proof. Assume that c is the smallest positive integer for which equation (12.2) holds. Then $(a, b) = 1$; indeed, if p is a prime such that $p \mid (a, b)$ then $p^4 \mid c^2$, so $p^2 \mid c$ by two applications of Euclid's Lemma Lemma 5.1[12.10] So a^2, b^2 and c form a primimite Pythagorean triple; hence, we may assume that a is odd. Thus, by Lemma 12.3, there are relatively primie positive integers x and y such that that $a^2 = x^2 - y^2$, $b^2 = 2xy$, and $c = x^2 + y^2$. Hence $a^2 = (x - y)(x + y)$. We have $(x - y, x + y) = 1$, since if p is a prime that $p \mid x - y$ and $p \mid x + y$ then $p \mid a$, and so p is odd, since a is odd. Further, $p \mid 2x = (x-y)+(x+y)$, and so $p \mid x$. Similarly $p \mid 2y = (x+y)-(x-y) = 2y$, and so $p \mid y$. This is, however a contradiction since x and y are relatively prime. Similarly as in the proof of Lemma 12.3, the equation $a^2 = (x - y)(x + y)$ shows that each prime factor of $x - y$ must occur with an even exponent; similarly for $x + y$. Thus, there are relatively prime positive integers u and v such that $x + y = u^2$ and $x - y = v^2$; here u and v are odd, since $x + y$ and $x - y$ are odd. Therefore, we have $2x = u^2 + v^2$ and $2y = u^2 - v^2$.

$$\frac{b^2}{4} = \frac{xy}{2} = \frac{u^2+v^2}{2}\frac{u^2-v^2}{4} = \frac{u^2+v^2}{2}\frac{u+v}{2}\frac{u-v}{2} = \alpha\beta\gamma,$$
$$\text{where} \quad \alpha = \frac{u^2+v^2}{2}, \quad \beta = \frac{u+v}{2}, \quad \text{and} \quad \gamma = \frac{u-v}{2}.$$

Here α, β, and γ are integers, since u and v are odd. Given that u and v are relatively prime, it follows that β and γ are relatively prime, by an argument similar to the one used twice before., We also have $u^2 = \alpha + 2\beta\gamma$ and $v^2 = \alpha - 2\beta\gamma$, by the same argument, it follows that $(\alpha, \beta\gamma) = 1$. Hence the same prime p can divide at most one the three numbers α, β, γ. Since $b^2/4 = (b/2)^2$ is the square of an integer, it follows by an argument we used before that there are positvie integers t, s, an r such that $\alpha = t^2$, $\beta = s^2$, and $\gamma = r^2$. Hence,

$$t^2 = \alpha = \frac{u^2+v^2}{2} = \left(\frac{u+v}{2}\right)^2 + \left(\frac{u-v}{2}\right)^2 = \beta^2 + \gamma^2 = s^4 + r^4.$$

Since $t^2 = \alpha = (u^2 + v^2)/2 = x < x^2 + y^2 = c$, this equation contradicts the minimality of c. □

What is interesting in this proof that the kind of induction we use supports the proof that the equation described in (12.2) cannot hold for integrs; that is, starting with the equation $a^4 + b^4 = c^2$ we can produce a similar equation with a smaller value of c. This argument would not support the proof of the weaker assertion that the equation $a^4 + b^4 = c^4$ never holds for positive integer values of a, b, and c, since starting with the latter equation, we cannot produce a similar equation with a smaller value of c; we can only produce an equation of the form $a^4 + b^4 = c^2$ with smaller integers, and not one of form $a^4 + b^4 = c^4$.

This means that it is easier to prove the stronger statement that the equation $a^4+b^4 = c^2$ cannot hold with positive integers instead of directly showing that the weaker statement the the equation

[12.9] French for "infinite descent."

[12.10] The first application shows that $c = kp$ for some integer k, but then we have $p \mid k^2$, so $p \mid k$ by the second application.

$a^4 + b^4 = c^4$ cannot hold for positive integers. The induction argument used only works with the former equation. That is, in this case it is easier to prove a more powerful result than the result we originally intended to prove (i.e., Fermat's Last Theorem for $n = 4$).

12.3 Reading

[2, Chapter 4, pp. 105–126].

12.4 Homework

[2, Chapter 4, pp. 105–126]. p. 110, 4.5, 4.7, 4.13.

13 Congruences

Definition 13.1. Let a, b, and c be integers. We say that a is *congruent*[13.1] to b modulo c, in symbols

$$a \equiv b \mod c,$$

if $c \mid b - a$.

13.1 Cancelation lemma

Lemma 13.1 (Cancelation lemma for congruences)**.** *Assume that*

$$ac \equiv bc \mod p,$$

where a, b, c are integers, p is a prime and $p \nmid c$. Then

$$a \equiv b \mod p.$$

Note that without the assumption that p is a prime, the conclusion might not be true. For example, we have

$$10 \cdot 2 \equiv 19 \cdot 2 \mod 6,$$

and yet

$$10 \not\equiv 19 \mod 6.$$

Proof. The congruence

$$ac \equiv bc \mod p$$

means that

$$p \mid bc - ac = (b - a)c.$$

A prime number is a divisor of a product only if it is a divisor of (at least) one of the factors; see Lemma 5.1. That is, for this to be true, we must have either $p \mid b - a$ or $p \mid c$ (recall that we assumed that p is a prime). The latter does not hold by our assumptions; so we must have $p \mid b - a$. That is,

$$a \equiv b \mod p.$$

This is what we wanted to show. □

The assumption $p \mid c$ can also be written as $c \not\equiv 0 \mod p$. Thus, the above lemma is analogous to the cancelation law for real numbers: if $ac = bc$ and $c \neq 0$ then $a = b$.

[13.1] The Latin word "modulo" is the ablative of the word modulus, which itself is the diminutive of "modus," meaning "measure." So "modulo" means "in a small measure."

13.2 Adding and multiplying congruences

Lemma 13.2. *Let a, b, c, d, and e be integers and assume that $a \equiv b \mod c$ and $d \equiv e \mod c$. Then $a + d \equiv b + e \mod c$ and $ad \equiv be \mod c$.*

Proof. The first congruene means $c \mid b - a$, hnd the second one, $c \mid e - d$. Hence $c \mid (b + e) - (a + d)$, establishing the third congruence. Further, $c \mid (b - a)d + b(e - d) = be - ad$, verifying the fourth congruence. □

13.3 Fermat's little theorem

Theorem 13.1 (Fermat)**.** *Let a be an integer and let p be a prime such that $p \nmid a$. Then $p \mid a^{p-1} - 1$.*

This is sometimes called Fermat's Little Theorem, as opposed to Fermat's Last Theorem 6.1.

Proof. For k with $1 \leq k \leq p - 1$, let a_k with $1 \leq a_k \leq p$ be such that[13.2]

$$a_k \equiv ak \mod p. \tag{13.1}$$

In other words, a_k is the remainder when ak is divided by p. For $k \neq l$ we have $a_k \neq a_l$, since otherwise we would have $a_k \equiv a_l \mod p$, i.e., $ak \equiv al \mod p$, but then the Cancelation Lemma 13.1 would imply that $k \equiv l$, which would mean that $k = l$. That is, the list of numbers a_1, a_2, ... a_{p-1}, is just a rearrangement of the list of numbers 1, 2, ..., $p - 1$. Hence

$$\prod_{k=1}^{p-1} a_k = \prod_{k=1}^{p-1} k. \tag{13.2}$$

Multiplying the congruences (13.1) together, we obtain

$$\prod_{k=1}^{p-1} a_k \equiv \prod_{k=1}^{p-1} ak \mod p.$$

By (13.2) this means that

$$1 \cdot \prod_{k=1}^{p-1} k \equiv a^{p-1} \prod_{k=1}^{p-1} k \mod p,$$

where we indicated multiplication by 1 so we can apply the Cancelation Lemma 13.1. As $p \nmid \prod_{k=1}^{p-1} k$ by repeated application of Euclid's Lemma 5.1, we obtain

$$1 \equiv a^{p-1} \mod p,$$

which is equivalent to what we wanted to prove. □

13.4 Residue classes

Given an integer c, the relation $a \equiv b \mod c$ on $\mathbb{Z}$ is reflexive, symmetric, and transitive, and so it is an equivalence relation; see Section 11. Thus we can define a partition of $\mathbb{Z}$ by putting

$$[n] = \{k \in \mathbb{Z} : n \equiv k \mod c\}$$

for $n \in \mathbb{Z}$; cf. Theorem 11.1. The sets $[n]$ are called residue classes modulo n. If $c > 0$, then the residue classes modulo c are [0], [1], [2], ..., $[c - 1]$, This is because if $n \in \mathbb{Z}$, there is a $k \in \mathbb{Z}$ with $0 \leq k < n$, and then $[n] = [k]$.

[13.2] The equality $a_k = p$ is clearly not possible, since ak is not divisible by p for $1 \leq k \leq p - 1$. We allowed $a_k = p$ because every number x is congruent modulo p to a number y with $1 \leq y \leq p$.

13.4.1 Compatibility with addition and multiplication

Let $f : S \times S \to S$ be a function; see Subsection 11.1. For convenience, one usually writes $f(x, y)$ for $x, y \in A$ instead of $f\big((x, y)\big)$, and calls f a function of two variables on S. Sometimes one prefers to write f as a binary operator; thus, instead of writing $f(x, y)$, one writes $x \diamond y$.

There is good reason for this in algebra. For example if one writes $A(x, y) = x + y$ for addition, and $M(x, y) = xy$ for multiplication, the distributivity of multiplication with respect to addition, expressed by the simple formula $a(b + c) = ab + ac$ becomes difficult to comprehend:

$$M\big(a, A(b, c)\big) = A\big(M(a, b), M(a, c)\big).$$

Definition 13.2. The binary operator $\diamond$ on the set A is said to be compatible with the equivalence relation R on A if, given $x, y, u, v \in A$, if xRu and yRv, then $(x \diamond u)R(y \diamond v)$.

If $\diamond$ is a binary operator compatible with the equivalence relation R on A, then one can extend the operation $\diamond$ to the equivalence classes $[x]_R$. Given two equivalence classes X and Y, by picking $x \in X$ and $y \in Y$, one can put

$$X \diamond Y \stackrel{def}{=} [x \diamond y]_R.$$

It is easy to show from the definition of compatibility that this definition is sound in that the right-hand side does not depend on the particular representatives x and y picked.

13.4.2 Operations on residue classes

Lemma 13.2 says that addition and multiplication is compatible with congruences. Thus, for the residue classes modulo c described above, we have can define addition and multiplication by stipulating $[a] + [b] = [a + b]$ and $[a][b] = [ab]$.

Problem 13.1. Let n be an integer. Show that $n^2 + 1$ is not divisible by 7.

Solution. The easiest approach is put the proof in terms of congruences. If $n \equiv 0 \mod 7$ then $n^2 + 1 \equiv 0^2 + 1 \equiv 1 \mod 7$. If $n \equiv \pm 1 \mod 7$ then $n^2 + 1 \equiv (\pm 1)^2 + 1 \equiv 2 \mod 7$. If $n \equiv \pm 2 \mod 7$ then $n^2 + 1 \equiv (\pm 2)^2 + 1 \equiv 5 \mod 7$. If $n \equiv \pm 3 \mod 7$ then $n^2 + 1 \equiv (\pm 3)^2 + 1 \equiv 10 \equiv 3 \mod 7$. There are no more possibilities. It is clear that in none of these cases does $n^2 + 1 \equiv 0 \mod 7$ holds. The proof is complete.[13.3]

Problem 13.2. Let $n \geq 0$ be an integer. Show that $3^n + 1$ is not divisible by 8.

Solution. We have $3^2 \equiv 1 \mod 8$; raising this the the power k, where $k \geq 0$ is an integer, we obtain $3^{2k} \equiv 1 \mod 8$. Multiplying this by 3, we can see that $3^{2k+1} \equiv 3 \mod 8$. Hence $3^n + 1 \equiv 2 \mod 8$ if n is even, and $3^n + 1 \equiv 4 \mod 8$ if n is odd.[13.4]

Problem 13.3. Let n be a positive integer. Show that

$$30^{10n+1} + 5^{21n}$$

is divisible by 31.

[13.3] See [11, Problem 2] for the source of the problem.
[13.4] See [14, Problem 4] for the source of the problem.

Solution. The proof of the assertion can perhaps be described in the language of congruences the simplest. For any integers $m > 0$, a, b, and c we have

$$ac + b \equiv b \mod c;$$

therefore, we have

$$(ac + b)^m \equiv b^m \mod c$$

by repeated aplications of Lemma 13.2. Using this, we have

$$\begin{aligned} 30^{10n+1} + 5^{21n} &= (31-1)^{10n+1} + (5^3)^{7n} = (31-1)^{10n+1} + (4 \cdot 31 + 1)^{7n} \\ &\equiv (-1)^{10n+1} + 1^{7n} \equiv -1 + 1 \equiv 0 \mod 31, \end{aligned}$$

establishing the assertion. [13.5][13.6]

13.5 Reading

[2, Chapter 4, pp. 110–112].

13.6 Homework

[2, Chapter 4, pp. 110–112]. p. 112, 4.19, 4.23.

14 Irrationality of square roots

Definition 14.1. A number $r \in \mathbb{R}$ is called *rational* if there are integers m and n such that $r = m/n$. A real number that is not rational is called *irrational*. The set of rational numbers is traditionally denoted by $\mathbb{Q}$.

14.1 The traditional proof of the irrationality of $\sqrt{2}$

Theorem 14.1. *The number $\sqrt{2}$ is irrational.*

Proof. Assume, on the contrary, that $\sqrt{2}$ is rational. Then there are integers m and n such that $\sqrt{2} = m/n$. We may assume here that m and n are positive and the fraction m/n is irreducible. After squaring the above equation, we obtain that $2 = (m/n)^2$, i.e., that $2 = m^2/n^2$. Multiplying both sides by n^2, we obtain that

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

The left-hand side here is even, so the right-hand side must also be even, i.e., m^2 is even. This means that m must be even, since the square of an odd number is always odd. Thus, we have $m = 2k$ for some integer k. Then $m^2 = (2k)^2 = 4k^2$, and so the above equation can be written as

$$2n^2 = 4k^2;$$

after dividing both sides by 2, we obtain that

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

[13.5] Note that in the last displayed line, the last two $\equiv$ symbols could be replaced with $=$; however, we did not want to surround the symbol $\equiv$ with equality symbols. So, after the first use of the symbol $\equiv$, we continued to use the symbol $\equiv$.

[13.6] See [17, Problem 4] for the source of the problem.

The right-hand side here is even, and so the left side must also be even; i.e., n^2 is even. This means that n must be even, since, as we mentioned just before, the square of an odd number is always odd.[14.1]

We obtained that both m and n even. Thus the fraction m/n can be reduced (by dividing both the numerator and denominator by 2, contradicting our assumption that the fraction m/n is irreducible. This assumption was based on our main assumption that $\sqrt{2}$ is rational. Since this latter assumption led to a contradiction, it must be false. Thus $\sqrt{2}$ must be irrational, as we wanted to show. □

14.2 Newer proofs of irrationality

Theorem 14.2. *Let n be a positive integer such that $\sqrt{n}$ is not an integer. Then $\sqrt{n}$ is irrational.*

The following proofs do not rely on the prime factorization of n. They are based on proofs that appeared at various places in the twentieth century. The first such place appears to be Carl B. Boyer's book on the history of mathematics, where the irrationality of $\sqrt{3}$ is proved. Later, the irrationality of $\sqrt{2}$ is proved along these lines by Theodor Estermann. Finally, Colin Richard Hughes, who was aware of Estermann's publication, used the method to prove the above result in its full generality. There is, however, no reason to assume that Boyer was not aware that the method is usable to prove the general result.[14.2]

First Proof. Assume $\sqrt{n}$ is rational. Let l be the smallest positive integer such that $\sqrt{n} = k/l$ for some integer k. Then $l\sqrt{n} = k$ and $k\sqrt{n} = (l\sqrt{n})\sqrt{n} = ln$. Let q be an integer such that $q < \sqrt{n} < q + 1$. Then

$$\sqrt{n} = \frac{k}{l} = \frac{k(\sqrt{n} - q)}{l(\sqrt{n} - q)} = \frac{k\sqrt{n} - kq}{l\sqrt{n} - lq} = \frac{ln - kq}{k - lq}.$$

On the right-hand side $\sqrt{n}$ is represented as the ratio of two integers. Since $k - lq = l(\sqrt{n} - q)$, and $0 < \sqrt{n} - q < 1$, we have

$$0 < k - lq < l,$$

which contradicts the choice of l, according to which l is the least positive integer for which $\sqrt{n} = k/l$. This contradiction shows that $\sqrt{n}$ is irrational. □

The second proof is essentially the same as the first proof, but it is explained somewhat differently.

Second Proof. Assume, on the contrary, that $\sqrt{n}$ is rational. Then there are positive integers k and l such that $\sqrt{n} = k/l$, i.e., such that $l\sqrt{n} = k$. Assume l is the smallest integer for which an integer k satisfying this equation exists. Let q and r be integers such that $k = ql + r$ and $0 \le r < l$. We cannot have $r = 0$ here since $r = 0$ would mean that $lq = k$, which, together with the equation $l\sqrt{n} = k$ would mean that $q = \sqrt{n}$, whereas we assumed that $\sqrt{n}$ is not an integer. Then we have $k\sqrt{n} = (l\sqrt{n})\sqrt{n} = ln$, and so

$$r\sqrt{n} = (k - ql)\sqrt{n} = k\sqrt{n} - ql\sqrt{n} = ln - qk.$$

Hence $r\sqrt{n}$ is an integer. This is, however, a contradiction, since $0 < r < l$ and l is the smallest positive integer for which $l\sqrt{n}$ is an integer. This contradiction proves that $\sqrt{n}$ is not rational. □

[14.1] A use of Euclid's Lemma 5.1 is involved here. We have $2 \mid n^2 = n \cdot n$. So, noting that 2 is a prime number, by the lemma we have $2 \mid n$ or $2 \mid n$, i.e., $2 \mid n$.

[14.2] See [1, p. 387], [3], and [5]. Boyer also describes geometric proofs of the irrationality of $\sqrt{2}$ and $\sqrt{5}$ that do not rely on prime factorization, either; see pp. 80–81 in his book quoted above.

14.3 Reading

Part of [2, §5.2, pp. 135 bottom–137].

14.4 Homework

[2, §5.2, p. 138]. p. 138, 5.21, 5.23.

15 Incommensurablitity: the golden ratio

15.1 The intercept theorem

The intercept theorem of Thales says that if two parallel lines cut two intersection lines, the ratio of the intercepts is the same. In Figre 15.1, the lines AB and CD are parallel, and the theorem asserts that $OA/OC = OB/OC$. It is not known how Thales justified this theorem, but he must have given a proof he found satisfactory. In the spirit of the times, and the subsequent crisis caused by he discovery of irrational numbers, he may have argued as follosws (see Figure 15.2). Find a line segment such OD such that both the segments OC and OA are integer multiples of the line segment OE, and then fill in the area inside the triangle $\triangle ABO$ with the small triangles as shown. Then the result follows simply by counting triangles.

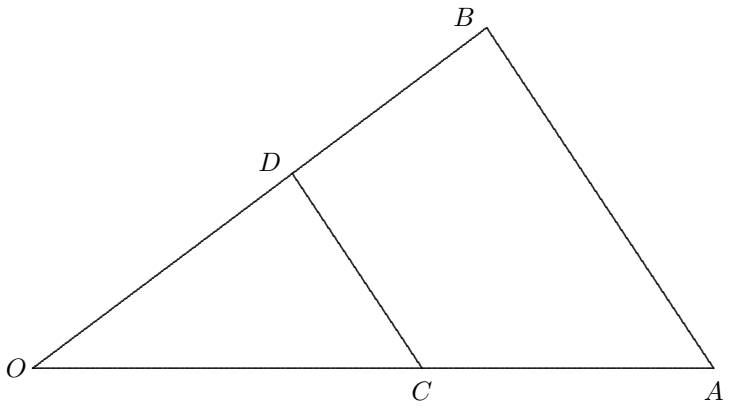

Figure 15.1: The intercept theorem

The ancient Greeks are known for their mathematical rigor, which was more demanding than in the two-hundred years after the discovery of calculus – when rigor was abandoned in favor of efficient discovery.[15.1] So the discovery that such a line segment may not exist precipitated a crisis in Greek mathematics. The discovery that such a line segment does not always exist is attributed to Hippasus, a member of the sect founded by Pythagoras. The Pythagoreans wanted to keep the discovery secret, to preemt a crisis in their mathematical thinking, so Hippasus's reward for revealing his discovery to the outside world was that he was drowned at sea. The crisis that this precipitated resolved by Eudoxus about two hundred years later.

[15.1] Only in the nineteenth century was calculus put on a rigorous foundation.

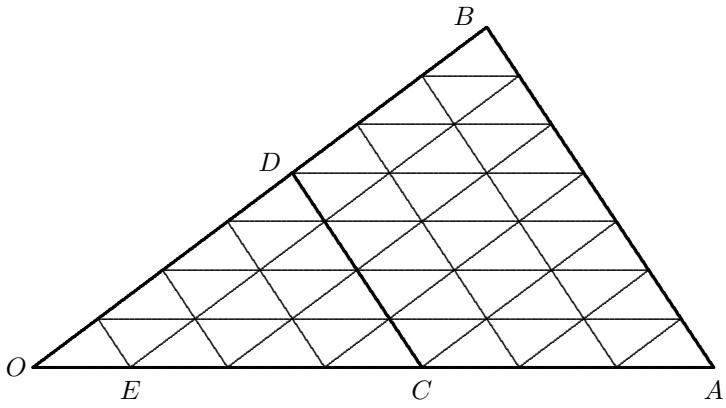

Figure 15.2: The intercept theorem explained

15.2 Commensurability

Definition 15.1. Let a and b be two real numbers. We call a and b *commensurable* if there is a real number $c \neq 0$ and there are integers k and l such that $a = kc$ and $b = lc$. If a and b are not commensurable, they are said to be *incommensurable*.

The ancient Greeks did not think in terms of numbers; they thought in terms of concrete quantities of their ratios; so, to follow their train of thought, we should have formulated the commensurablity of two quantities, such as two line segments (or areas). Assuming $b \neq 0$, it is easy to see that a and b are commensurable if and only if the ratio a/b is a rational number, i.e., it is the ratio of two integers. Indeed, if a and b are commensurable, then with the c above, we have

$$\frac{a}{b} = \frac{kc}{lc} = \frac{k}{l}.$$

On the other hand, if $a/b = k/l$ for some integers k and l, then

$$\frac{a+b}{b} = \frac{a}{b} + 1 = \frac{k}{l} + 1 = \frac{k+l}{l}.$$

Except in the case $a + b = 0$ (when $a = -b$, and so a and b are commesurable, as witnessed by the quantity $c = a$), this implies

$$b = l\,\frac{a+b}{k+l} \quad \text{and} \quad a = \frac{k}{l}\,b = \frac{k}{l}\,l\,\frac{a+b}{k+l} = k\,\frac{a+b}{k+l}.$$

Thus, the quantity $(a + b)/(k + l)$ witnesses the commensurability of a and b.

15.3 The regular pentagon

Figure 15.3 shows a regular pentagon. We wish to determine the ratio of the diagonal BE and the side AB. To this end, note that

$$\angle BAF = \angle CDA = \angle BFA;$$

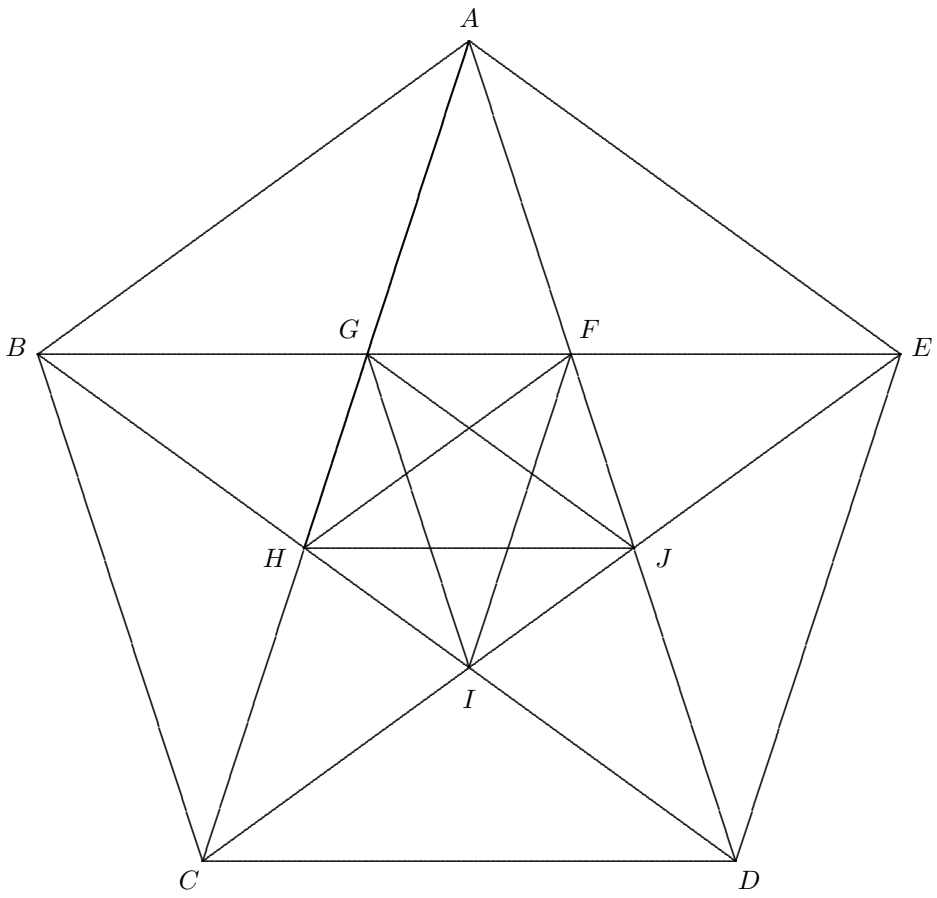

Figure 15.3: Regular pentagon

the first equation holds by symmetry, and the second equation holds becasue the lines CD and BF are parallel. The equality of the angles on the sides mean that the triangle $\triangle ABF$ is isosceles, i.e, $AB = FB$. For reasons of symmetry, $FB = GE$. Writing a_1 for the side of the pentagon $ABCDE$, d_1 for the diagonal of same pentagon, a_2 for the side of the smaller pentagon $FGHIJ$, and d_2 for its diagonal, we have $d_1 = BF + GE - GF = 2a_1 - a_2$; that is

$$a_2 = 2a_1 - d_1. \tag{15.1}$$

Further

$$\angle HBF = \angle JEG = \angle HFB;$$

again, the first equality holds by symmetry, and the second equation hods since the lines JE are parallel HF are parallel (both are parallel to BA). Thus, the triangle $\triangle BHF$ is isosceles, and $BH = HF$. For reasons of symmetry, $BH = BG$. That is, writing d_2 for the diagonal of the pentagon $FGHIJ$, we have

$$d_2 = HF = BG = BE - EG = d_1 - a_1. \tag{15.2}$$

That is

$$d_2 = d_1 - a_1. \tag{15.3}$$

Dividing this equation by a_1, we obtain that

$$\frac{d_2}{a_1} = \frac{d_1}{a_1} - 1.$$

Noting that $\triangle BGA \sim \triangle BAE$, it follows that $BG/BA = BA/BE$; since $BG = d_2$ according to equation (15.2), this can be written as $d_2/a_1 = a_1/d_1$, Hence, the above equation becomes

$$\frac{a_1}{d_1} = \frac{d_1}{a_1} - 1.$$

Multiplying this equation by d_1/a_1 and rearranging this equation, we obtain

$$\left(\frac{d_1}{a_1}\right)^2 - \frac{d_1}{a_1} - 1 = 0.$$

Using the quadratic formula to solve this equation for d_1/a_1, we obtain that

$$\frac{d_1}{a_1} = \frac{1 \pm \sqrt{1+4}}{2} = \frac{1 \pm \sqrt{5}}{2}.$$

Here the $-$ sign in the numerator gives a negative solution, so we must take the $+$ sign:

$$\frac{d_1}{a_1} = \frac{1 + \sqrt{5}}{2}. \tag{15.4}$$

This number is called the golden ratio.

15.4 Incommensurability

Since we already know that $\sqrt{5}$ is irrational, from equation (15.4) it follows that the golden ratio is irrational. In other words, the side and the diagonal of the regular pentagon are incommensurable. It will be interestig to establish this directly, especially since the famous historian of mathematics, Carl Boyer, who was a professor of mathematics at Brooklyn College, speculated in his book [1, p. 80, pdf p. 97] that incommensurable quantities were discovered by showing the incommensurability of the side of the side and diagonal of the regular pentagon.

In Figure 15.3 a large pentagon is shown, and smaller pentagons obtained by drawing the diagonals of this and the resulting pentagons are shown. While only three of them appear in the diagram, this process can be continued indefinitely. Relations (15.1) and (15.3) express the side and the diagonal of the second pentagon in terms of the side and the diagonal of the first pentagon. Writing a_n and d_n for the side and the diagonal of the nth pentagon, these relations can obviously extended to any of these pentagons and the next pentagon formed by its diagonals. Thus, we have

$$a_{n+1} = 2a_n - d_n \quad \text{and} \quad d_{n+1} = d_n - a_n \qquad (n \geq 1). \tag{15.5}$$

Now, assume that a_1 and d_1 are commensurable, i.e., that there is a line segment $c > 0$ and there are integers k_1 and l_1 such that $a_1 = k_1 c$ and $d_1 = l_1 c$. For $n > 1$, let the real numbers k_n and l_n be

defined as $k_n = a_n/c$ and $l_n = d_n/c$. Then $a_n = k_n c$ and $d_n = l_n c$; we will show that k_n and l_n are integers. Indeed, equations (15.5) imply that equation

$$k_{n+1}c = a_{n+1} = 2a_n - d_n = 2k_n c - l_n c = (2k_n - l_n)c$$

and

$$l_{n+1}c = d_{n+1} = d_n - a_n = l_n c - k_n c = (l_n - k_n)c.$$

Dividining these equations by c, we obtain

$$k_{n+1} = 2k_n - l_n \quad \text{and} \quad l_{n+1} = l_n - k_n.$$

It follows that if k_n and l_n are integers then k_{n+1} and l_{n+1} are also integers. Since k_1 and l_1 were assumed to be integers, it follows then that k_2 and l_2 are integers, and so on, continuing this way it follows that k_n and l_n are integers for all n. Since a_n is a line segment of positive length, it follows that $a_n = k_n c \geq c$ for all n. This is, however, impossible since a_n, the side of the nth pentagon can be made arbitrarily small; that is, there is an n for which $a_n < c$, which is a contradiction. This contradiction shows that k_1 and l_1 cannot be integers. In other words, the line segment a_1 and d_1 are incommensurable.

16 Continued fractions

Consider the following equations

$$\frac{546}{422} = 1 + \frac{124}{422} = 1 + \frac{1}{\frac{422}{124}} = 1 + \frac{1}{3 + \frac{50}{124}} = 1 + \frac{1}{3 + \frac{1}{\frac{124}{50}}} = 1 + \frac{1}{3 + \frac{1}{2 + \frac{24}{50}}}. \tag{16.1}$$

We will give a brief description what we do here. Given a number t, denote by $\lfloor t \rfloor$ its i*integer part*; i.e., $\lfloor t \rfloor$ is the largest integer n such that $n \leq t$, and write $\{t\} = t - \lfloor t \rfloor$, called the fractional part of t.[16.1] We start with the number $x = 546/422$, we take its fractional part $\{x\} = 124/422$; then we take reciprocal $y_1 = 1/\{x\} = 422/124$. Then we take its fractional part $\{y_1\} = 50/124$. Again, we take its reciprocal $y_2 = 1/y_1 = 124/50$. Taking fractional part again, we have $\{y_2\} = 24/50$. We will continue this by taking reciprocals again:

$$\begin{aligned} \frac{546}{422} &= 1 + \frac{1}{3 + \frac{1}{2 + \frac{1}{\frac{50}{24}}}} = 1 + \frac{1}{3 + \frac{1}{2 + \frac{1}{2 + \frac{2}{24}}}} = 1 + \frac{1}{3 + \frac{1}{2 + \frac{1}{2 + \frac{1}{\frac{24}{2}}}}} \\ &= 1 + \frac{1}{3 + \frac{1}{2 + \frac{1}{2 + \frac{1}{12}}}}. \end{aligned} \tag{16.2}$$

[16.1] Traditionally, the integer part of t used to be denoted as $[t]$; of course, this notation can also have other meanings, and with computerized typesetting the more specific notation $\lfloor t \rfloor$ was introduced. For the fractional part, for lack of a better notation, the traditional notation $\{t\}$ is retained.

The fact that the starting fraction 546/422 is reducible has no importance here; in any case, it does not influence the final result. The expression on the wrine hand side is called the continued fraction representation of 546/422. In order to save space, one may prefer to write the fraction on the right-hand side as

$$1+\frac{1}{3+\frac{1}{2+\frac{1}{2+\frac{1}{12}}}};$$

note the sign + placed below the main line to indicate the the addition happens in the denominator of the preceding fraction. Here the numbers

$$1,\quad 1+\frac{1}{3},\quad 1+\frac{1}{3+\frac{1}{2}},\quad 1+\frac{1}{3+\frac{1}{2+\frac{1}{2}}},\quad 1+\frac{1}{3+\frac{1}{2+\frac{1}{2+\frac{1}{12}}}}$$

are called the continued fraction approximants of the number we started with. These are in fact the expressions obtained by omitting the last fraction in the 2nd, 4th, 6th members of the equations in (16.1) and in the 3rd member of equations in (16.2) and and its last (5th) member, in which the last fraction is not omitted).

It is easy to describe in general terms how these approximants are obtained. Let x denote the number to be apporimated (546/422 at present). Given a number t, denote by $\lfloor t\rfloor$ its i*nteger part* and by $\{t\}=t-\lfloor t\rfloor$ its fractional part; note that for any t we have $0\leq\{t\}<1$. Starting with x, we write $x=\lfloor x\rfloor+\{x\}$. We have $x_0=\lfloor x\rfloor$ the first approximant. If $\{x\}=0$ then we are finished. If $\{x\}\neq 0$, we write $y_1=1/\{x\}$; then $y_1\geq 1$, and we write $y_1=\lfloor y_1\rfloor+\{y_1\}$. We take

$$x_1=x_0+\frac{1}{\lfloor y_1\rfloor}.$$

If $\{y_1\}=0$, we are finished. if $\{y_1\}\neq 0$, $y_1=\lfloor y_1\rfloor+\{y_1\}$. we take $y_2=1/\{y_1\}$, and write $y_2=\lfloor y_2\rfloor+\{y_2\}$.

$$x_2=x_0+\frac{1}{\lfloor y_1\rfloor}+\frac{1}{\lfloor y_2\rfloor}.$$

Coninuing this way, if $y_k\geq 1$ has been defined, we write

$$x_k=x_0+\frac{1}{\lfloor y_1\rfloor}+\frac{1}{\lfloor y_2\rfloor}+\cdots+\frac{1}{\lfloor y_k\rfloor}.$$

If $\{y_k\}\neq 0$, and we write $y_k=\lfloor y_k\rfloor+\{y_k\}$. and we take $y_{k+1}=1/\{y_k\}$. If $\{y_k\}=0$, then the process is finished and we have $x=x_k$.

16.1 The irrationality of the golden ratio revisited

Note that the starting number x may be negative, but all the numbers y_k will be positive except for the last one (in case the process ends in a finite number of steps. If we start with a rational number x, then it is clear that the process ends in a finite number of steps; this is because all the numbers y_k are fractions, and the numerators and denominators of these fractions decrease. In fact, in the introductory example presented in formula (16.1) we have $x=546/422$, $y_1=422/124$, $y_2=124/50$, $y_3=50/24$, $y_4=24/2=12$, and y_5 is not defined, since y_4 is an integer. So, if we start with a number x and we are led to an infinite continued fraction, then this shows that x is irrational. This gives a simple proof that the golden ratio $(1+\sqrt{5})/2$ given in equation (15.4) is an irrational

number. Indeed, we have

$$\frac{1+\sqrt{5}}{2} = 1 + \frac{\sqrt{5}-1}{2} = 1 + \frac{(\sqrt{5}-1)(\sqrt{5}+1)}{2(\sqrt{5}+1)} = 1 + \frac{4}{2(\sqrt{5}+1)} = 1 + \frac{2}{\sqrt{5}+1}$$
$$= 1 + \frac{1}{\frac{1+\sqrt{5}}{2}}$$

In the last denominator we ended up with the same golden ratio that we started with. So, when we continue the process, things will just repeat; for example

$$\frac{1+\sqrt{5}}{2} = 1 + \frac{1}{1+\frac{1}{\frac{1+\sqrt{5}}{2}}} = 1 + \frac{1}{1+\frac{1}{1+\frac{1+\sqrt{5}}{2}}} = 1 + \frac{1}{1+\frac{1}{1+\frac{1}{1+\frac{1+\sqrt{5}}{2}}}}$$
$$= 1 + \frac{1}{1} + \frac{1}{1} + \frac{1}{1} + \cdots.$$

Given that this fraction goes on indefinitely, it follows that $(1+\sqrt{5})/2$ is irrational.

It is not difficult to show that this proof is just a reformulation of the proof that the diagonal and the side of the regular pentagon are incommensurable, given in Subsection 15.4.

16.2 The case of $\sqrt{2}$

Other square roots can be shown to be irrational in a similar way. For example, we have

$$\sqrt{2} = 1 + (\sqrt{2}-1) = 1 + \frac{(\sqrt{2}+1)(\sqrt{2}-1)}{\sqrt{2}+1} = 1 + \frac{1}{\sqrt{2}+1} = 1 + \frac{1}{2+(\sqrt{2}-1)}$$

Since the expresssion $\sqrt{2}-1$ occurs the second time, the process will repeat when continuing it. We have

$$\sqrt{2} = 1 + \frac{1}{2+\frac{1}{2+(\sqrt{2}-1)}} = 1 + \frac{1}{2+\frac{1}{2+\frac{1}{2+(\sqrt{2}-1)}}} = 1 + \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \cdots.$$

This this continued fraction goes on indefinitely, it follows that $\sqrt{2}$ is irrational.

16.3 Other square roots

Assume $n > 0$ is and integer whose square root is not an integer, and let q be an integer such that $q < \sqrt{n} < q+1$. If one tried to get the continued fraction expansion of $\sqrt{n}$, first writes it as $\sqrt{n} = \lfloor\sqrt{n}\rfloor + \{\sqrt{n}\}$, where, as before $\lfloor x\rfloor$ denotes the integer part of x, and $\{x\}$ denotes its fractional part. That is,

$$\sqrt{n} = q + (\sqrt{n}-q) = q + \frac{(\sqrt{n}-q)(\sqrt{n}+q)}{\sqrt{n}+q}$$
$$= q + \frac{n-q^2}{2q+(\sqrt{n}-q)} \tag{16.3}$$

A part of this caclulation show that

$$\sqrt{n} - q = \frac{n - q^2}{2q + (\sqrt{n} - q)} \tag{16.4}$$

Using this equation repeatedly for the to continue the expansion started in equation (16.3), we can see that

$$\begin{aligned}\sqrt{n} = q + (\sqrt{n} - q) &= q + \cfrac{n - q^2}{2q + \cfrac{n - q^2}{2q + (\sqrt{n} - q)}} \\ &= q + \frac{n - q^2}{2q} + \frac{n - q^2}{2q} + \frac{n - q^2}{2q} + \frac{n - q^2}{2q} \cdots = q + \operatorname*{\mathcal{K}}_{k=1}^{\infty} \frac{n - q^2}{21};\end{aligned}$$

the right-hand side uses Gauss's notation for continued fraction, in analogy for with the the sigma notation for sums. The problem with this expansion is that it is a generalized form of the expansion given in equation (16.2); in that expansion, the numeratons are all 1. In a regular continued fraction expansion, it is required to be that all the 1, where this is not required, the expansion is called a *generalized continued fraction expansion*. The infinity is a generalized contiued fraction expansion does not guarantee the irrationality of the number so represented, without additional conditions. On the other hand, equation (16.4) in itself inspires another proof of irrationality of $\sqrt{n}$ is $\sqrt{n}$ is not an integer.

To see how this is done, assume we will use the following lemma:

Lemma 16.1. *Assume k, l, q, and r are integers such that $(k, l) = 1$ and*

$$\frac{k}{l} = \frac{q}{r}.$$

Then $k \mid q$ and $l \mid r$.

Proof. We have $kr = lq$, so $k \mid lq$. Since $(k, l) = 1$, $k \mid q$ follows by Corollary 5.2. Similarly, we have $l \mid kr$ and so $l \mid r$ follows by the same lemma. □

Now, assuming $\sqrt{n} - q = k/l$, where $(k, l) = 1$, by equation (16.4) we have

$$\frac{k}{l} = \frac{n - q^2}{2q + \frac{k}{l}} = \frac{l(n - q^2)}{2ql + k}.$$

Since the fractions on the left-hand side and the right-hand side are equal, Lemma 16.1 implies that $l \mid 2ql + k$. Hence $l \mid k$, which is imposssible, since we assumed that $\sqrt{n}$ is not an integer. Hence $\sqrt{n} - q$ is irrational, and so $\sqrt{n}$ is also irrational.

It would be interesting to explore the connection of this proof to the proofs in Subsection 14.2.

16.4 The irrationality of π

In 1761, Johann Heinrich Lambert proved that π is irrational. He used the infinity of certain generalized continued fractions, but in a way too complicated for inclusion in these notes. For the interested reader, the proof can be found online at Lambert's proof. The proof is mainly written in English, with occasional short French phrases within formulas (presumably because these were pasted from the French version); with available online dictionaries such as such as Wiktionary, the French phrases can easily be translated even by those with no prior knowledge of French.

17 The greatest common divisor

Lemma 17.1 (Division Theorem)**.** *Let a and d be integers, $d \neq 0$. Then there are integers r and q such that*

$$a = qd + r \qquad \text{and} \qquad 0 \leq r < |d|.$$

The book [2, p. 306], and many other sources, call this result *Division Algorithm*, while we call it Division Theorem (even though we state it as a lemma). The reason is that algorithm means a method of calculation, usually described in a systematic way as to the steps to be followed, whereas a proven mathematical statement is usually called a theorem, a lemma, or something similar (such as an observation in case the proof is very simple). In the result, a is called the dividend, d, the divisor, q, the quotient, and r, the remainder. In the actual division algorithm taught in schools, given a and d, one uses the method to determine d and r. Since this result is so simple, it is often used without proof, and in fact without even mentioning explicitly that the result was used. In fact, it was used in the proofs of Lemmas 5.1 (Euclid's lemma) and 5.2 (the key lemma used in the second proof of Euclid's lemma).

Proof. Consider the set

$$S = \{a - dx \geq 0 : x \in \mathbb{Z}\}.$$

This set is clearly not empty, as witnessed by the choice $x = -da^2$. iLet $r = a - dx_0$ be the smallest element of this set. Then $0 \leq r < |d|$ (the first inequality follows from the definition of S). Indeed, assuming $r \geq |d|$, the number $r' = a - d(x_0 + 1)$ (in case $d > 0$) or $r'' = a - d(x_0 - 1)$ (in case $d < 0$) is a smaller element of this set, a contradiction. The lemma is then satisfied with this r and $q = x_0$. □

Definition 17.1. Let a and b be integers. The greatest common divisor d of a and b is defined as the largest integer d such that $d \mid a$ and $d \mid b$. The greatest common divisor of a and b is often written as (a, b); here we will also use the notation $\gcd(a, b)$.

Note that any number is a divisor of 0; hence $\gcd(0, 0)$ is not defined. If a and b are integers not both of which are 0 then $\gcd(a, b)$ is defined; even if a or b are negative, $\gcd(a, b) > 0$; this is simply because any positive integer is larger than any negative integer. The key lemma in studying the greatest common divisor is the following:

Lemma 17.2. *Let a and b be integers, not both of which are 0. Then there are integers x and y such that*

$$\gcd(a, b) = ax + by.$$

The identity in this lemma is called Bézout's identity. However, the identity was known to Euclid, since it is a direct consequence of the Euclidean algorithm.

Proof. Consider the set.

$$S = \{ax + by > 0 : x, y \in \mathbb{Z}\}.$$

This set is clearly not empty, as witnessed by the choice $x = a$ and $y = b$. Let d be its smallest element. First we will show that $d \mid a$. Assume, on the contrary that $d \nmid a$. Let q and r be such that $a = dq + r$ and $0 < r < d$. Then, noting that we have

$$d = ax_0 + by_0$$

for some integers x_0 and y_0, we have

$$r = a - dq = a(1 - qx_0) + b(-qy_0),$$

showing that $r \in S$. This is a contradiction, since d is the smallest element of S. Similarly, we can show that $d \mid b$.

Next, assume that $c > d$ is a common divisor of a and b. Then $a = cx_1$ and $b = cy_1$. Then

$$d = ax_0 + by_0 = c(x_1x_0) + c(y_1y_0) = c(x_1x_0 + y_1y_0).$$

Since $x_1x_0 + y_1y_0 \geq 1$, this shows that $d \geq c$, again a contradiction. □

Corollary 17.1. *Let a and b be integers and let p be a prime. If ab is divisible by p then either a is divisible by p or b is divisible by p.*

Note that this is just Euclid's lemma, stated above as Lemma 5.1; here we will give a new proof.

Proof. We will suppose that a and b are both positive (the case when a or b equals 0 can easily be dealt with, and when a or b is negative, it is harmless to consider their absolute values instead). Assume that a is not divisible by p; we will then have to show that b is divisible by p. With this assumption, we have $(a, p) = 1$. Indeed, p being a prime number, its only divisors are 1 and p. As we assumed that p is not a divisor of a, the greatest common divisor of a and p can only be one. As we saw above, this greatest common divisor can be represented as a linear combination. That is, we have integers x and y such that

$$1 = ax + py.$$

Multiplying both sides by b, we obtain

$$b = abx + pby.$$

Both terms on the right-hand side are divisible by p; indeed, our assumption was that ab is divisible by p. Hence the left-hand side, that is, b, is also divisible by p. This is what we wanted to show. □

Note that it is easy to conclude from the above lemma a similar statement for the product of more than two integers. For example, if, for some integers, a, b, and c, and for some prime number p, the product abc is divisible by p, then either a or b or c is divisible by p. To see this, write $abc = (ab)c$. Then, by using the lemma, we can see that either ab or c is divisible by p. If now ab is divisible by p, then, using the lemma again, we can see that either a or b is divisible by p. Similarly for the product of four or more integers.

We can now use the above lemma, or, rather, its extension to a product of more than two factors, to give another proof of the Unique Factorization Theorem, already stated as Theorem 5.1

Corollary 17.2. *The prime factorization of every integer greater than one is unique, aside from the order of factors.*

Proof. Assume, on the contrary, that there are positive ,integers with two different prime factorizations. Let n be the smallest such integer, and let its two different prime factorizations be

$$n = p_1p_2 \dots p_k = q_1q_2 \dots q_l.$$

Note first that each of the primes p_i must be different from each of the primes q_j. Indeed, if, for example, we had $p_1 = q_1$ then the number

$$n/p_1 = p_2p_3 \dots p_k = q_2q_3 \dots q_l$$

would be a number smaller than n with two different prime factorizations, in contradiction with our assumption.

On the other hand, the equation above shows that the prime number p_1 is a divisor of the product $q_1 q_2 \dots q_l$. So, by the lemma above, or, rather, by its extension to more than two factors, at least one of the numbers q_1, q_2,... q_l, say q_j for a certain j with $1 \leq j \leq l$, must be divisible by p_1. As q_j is a prime, its only divisors are 1 and itself. As p_1, being a prime, is different from 1, we must have $p_1 = q_j$. This is a contradiction, since, as we stated above, p_1 must be different from q_j. This contradiction shows that our initial assumption was wrong; that is, there are no integers with two different prime factorizations. □

17.1 Reading

[2, Chapter 12, pp. 303–326].

17.2 Homework

[2, 5.2, p.138], p. 138, 5.25, [2, Chapter 12, pp. 303–326], p. 305, 12.1, 12.3, 12,7, p. 309, 12.19, 12.27, p. 312, 12.35, p. 315, 12.27, 12.41, p. 317, 12.51, 12.53, 12.55, p. 321, 12.67, p. 323, 12.71.

18 Functions

Definition 18.1. A *function* is a set of ordered pairs where the first member determines the second member.

Formally,

$$\text{Function}(f) \leftrightarrow \forall x \in f\, \exists y\, \exists z \big(x = (y,z)\big)\, \& \forall x \forall y \forall z \Big(\big((x,y) \in f\, \& (x,z) \in f\big) \rightarrow y = z \Big). \tag{18.1}$$

Note the reuse of the variable x in the formula to indicate comletely different things. To clarify the meaning of this formula, one often introduces the quantifier $\exists!$ to mean there exist exactly one. That is, if the variable y is not free in $A(x)$, then $\exists! x A(x)$ is an abbreviation of

$$\exists x \Big(\phi(x)\, \& \forall y \big(\phi(y) \rightarrow x = y\big) \Big).$$

With this abbreviation, we can also write

$$\text{Function}(f) \leftrightarrow \forall x \in f\, \exists y\, \exists z \big(x = (y,z)\big)\, \& \forall x \Big(\exists y \big((x,y) \in f\big) \rightarrow \exists! y \big((x,y) \in f\big) \Big).$$

Admittedly, this is not really simpler than the first version, but the description given by this formula is closer to the verbal description given above.

18.1 Domain, range, inverse

The domain and the range of a function is defined as in case of relations in Subsection 11.2:

$$\text{dom}(f) \stackrel{def}{=} \{x : \exists y\, (x,y) \in f\}, \tag{18.2}$$

and

$$\text{ra}(f) \stackrel{def}{=} \{y : \exists x\, (x,y) \in f\}. \tag{18.3}$$

Givn $x \in \mathrm{dom}(f)$, the *value* $f(x)$ of f at x is defined as the unique u for which $(x, y) \in f$:

$$y = f(x) \leftrightarrow (x, y) \in f. \tag{18.4}$$

The inverse of a function can be defined as a relation that results by reversing the pairs:

$$f^{-1} \stackrel{def}{=} \{(x, y) : (y, x) \in f\}; \tag{18.5}$$

in case f^{-1} is a function, we call it the *inverse function*, or more simply, the *inverse*, of f. If it is not a function, we will continue to use the phrase "inverse as a relation." Given that f^{-1} is not necessarily a function, we need a further discussion, given below.

18.2 Function from a set into another

The symbol $f : A \to B$ indicates that f is a function such that $\mathrm{dom}(f) = A$ and $\mathrm{ra}(f) \subset B$; one also says that f is a function from A into B. If one does not want to mention B at all, one says f is a function on A; this means that $\mathrm{dom}(f) = A$. This contrast with the usage for relations, since in case of relations, R being on a relation on the set A means that $\mathrm{dom}(R) \subset A$ and $\mathrm{ra}(R) \subset A$. To express that $\mathrm{ra}(f) = B$, one says that f is a function onto B.

The set B is often called the *codomain* of f, though one should be careful to use this term, since altering the codomain does not alter the function. For example, writing $C = \{x \in \mathbb{R} : x \geq 0\}$, if we define the function f as $f : \mathbb{R} \to \mathbb{R}$ such that $f(x) = x^2$ for all $x \in \mathbb{R}$, and the function g as $g : \mathbb{R} \to C$ such that $g(x) = x^2$ for all $x \in \mathbb{R}$, then f and g are the same function, that is, we have $f = g = \{(x, x^2) : x \in \mathbb{R}\}$. This is true notwithstanding the fact that the codomain of f was specified as $\mathbb{R}$ and that of g as C.

There is a good reason to specify a codomain instead of the range when one describes a function: when describing the function, finding a suitable codomain is usually a simple matter, but finding the range may involve a substantial effort, and occasionally it cannot be done exactly. For eample, one can satisfactorily describe a function f by saying that its domain is the interval $[0, 2]$, and for $x \in [0, 2]$ $f(x) = x^6 - 3x^2 - x$, but to determine its range requires a substantial effort. If one says $f : [0, 2] \to \mathbb{R}$, one does not need to specify the range.

18.3 Composition

For functions, one can also define a new operation, called composition:[18.1]

Definition 18.2 (Composition of functions)**.** Let f and g be functions. Then the composition of f and g is defined as

$$f \circ g \stackrel{def}{=} \big\{(x, y) : (\exists z)\big((x, z) \in g \,\&\, (z, y) \in f\big\}\,. \tag{18.6}$$

There is a certain awkwardness in this definition, since given any two relations, R and S, it is more natural to define their composition as

$$R \circ S \stackrel{def}{=} \big\{(x, y) : (\exists z)\big((x, z) \in R \,\&\, (z, y) \in S\big\}\,.$$

[18.1] One can define composition also in case of relations, but we did not do so, since it was not important for our discussions. The definition is the same as we are about to give for functions.

Gödel [4] resolved this difficulty by defining the function f as the set of pairs $\big(f(x), x\big)$. In our discussions, f is the set of pairs $\big(x, f(x)\big)$.

Gödel's solution is very elegant; the only reason we did not adopt it is because it is not commonly used in the literature. Unfortunately, Gödel did not explain his reasons for the way he defined functions, so the beauty of his approach is not widely appreciated.

It is easy to see that if $x \in \mathrm{dom}(f \circ g)$ then $(f \circ g)(x) = f\big(g(x)\big)$, and this equation is often used to define the composition of the functions f and g. Our definition, however, simplifies the discussion of composition.

One can reformulate the definiton of composition as follows. Given functions f and g, their composition $f \circ g$ is a relation such that

$$(x, y) \in f \circ g \leftrightarrow (\exists z)\big((x, z) \in g \,\&(z, y) \in f\big). \tag{18.7}$$

The only difference here as compared to formula (18.6) is that there the formula expresses that $f \circ g$ is a set of ordered pairs; here we expressed that in words by saying that $f \circ g$ is a relation. In cases where it is clear that we are dealing with relations, it will be simpler to refer to formula (18.7) than to (18.6).

Lemma 18.1. *The composition of two functions is a function.*

Proof. Let f and g be functions. It is clear that $f \circ g$ is a relation, so what we need to show is that given any x, there is at most one y such that $(x, y) \in f \circ g$. Looking at the right-hand side of the biconditional in formula (18.7), given x, there is a most one z for which $(x, z) \in g$ since g is a function. If we can find such a unique z, again there is at most one y for which $(z, y) \in f$, since f is a function. Hence, given x, the right-hand side of the biconditional can be satisfied by at most one z. □

Lemma 18.2 (Associativity of composition)**.** *The composition of functions is associative.*

Proof. In the proof, we will use the symbol $\equiv$ in accordance with what we said in Subsection 3.3.[18.2] Given three functions f, g, and h, we need to show that $f \circ (g \circ h) = (f \circ g) \circ h$. We have

$$\begin{aligned}(x, y) \in f \circ (g \circ h) &\equiv (\exists z)\big((x, z) \in (g \circ h) \,\&(z, y) \in f\big)\\ &\equiv (\exists z)\Big((\exists u)\big((x, u) \in h \,\&(u, z) \in g\big) \,\&(z, y) \in f\Big)\\ &\equiv (\exists z)(\exists u)\big((x, u) \in h \,\&(u, z) \in g \,\&(z, y) \in f\big);\end{aligned}$$

In the last step, the part $\&(z, y) \in f$ of the formula was brought under the scope of the quantifier $(\exists u)$. This is permissible, since u does not occur free in this part.[18.3]

Similarly, we have

$$\begin{aligned}(x, y) \in (f \circ g) \circ h &\equiv (\exists u)\big((x, u) \in h \,\&(u, y) \in f \circ g\big)\\ &\equiv (\exists u)\Big((x, u) \in h \,\&(\exists z)\big((u, z) \in g \,\&(z, y) \in f\big)\Big)\\ &\equiv (\exists u)(\exists z)\big((x, u) \in h \,\&(u, z) \in g \,\&(z, y) \in f\big);\end{aligned}$$

In the last step, we brought the part $(x, u) \in h \,\&$ under the scope of the quantifiles $(\exists z)$. This is permissible, since z does not occur free in this part. Noting that adjacent existential quantifiers are interchangeable, the right-hand sides of both formulas are equivalent, establishing the result. □

[18.2] As long as we were comparing two logic expressions, there was no difference between using $\leftrightarrow$ and $\equiv$.

[18.3] In fact, u does not occur at all, but only free occurrences matter.

This result can, of course, be established verbally, without the use of logic formulas; this is done in [2, Theorem 10.13, p. 266]. But in such a proof, the words obscure the key facts of logic on which the proof is based. The reader is urged to read both proofs carefully to appreciate this.

18.3.1 Domain of a composition

According to formula (18.7) for x to be in the domain for $f \circ g$, we need to have $x \in \mathrm{dom}(g)$, and for $z = g(x)$ we must have $z \in \mathrm{dom}(f)$. That is,

$$\mathrm{dom}(f \circ g) = \{x \in \mathrm{dom}(g) : g(x) \in \mathrm{dom}(f)\}. \tag{18.8}$$

18.4 Inverse function

The inverse of a function as a relation was given in formula (18.5). One can restate this by saying that f^{-1} is a relation such that

$$(y, x) \in f^{-1} \leftrightarrow (x, y) \in f \tag{18.9}$$

for all x and y. We mentioned, however, that f^{-1} is not necessarily a function. The condition needed for f^{-1} to be a function is easy to describe: for the function f, if $(x, y) \in f$, then y must uniquely determe x. Such functions are called one-to-one. Formally, assuming f is a function,

$$\text{One-to-one}(f) \leftrightarrow (\forall y \in \mathrm{ra}(f))(\exists! x)\big((x, y) \in f\big). \tag{18.10}$$

One can also express this more directly, without mentioning the domaim:

$$\text{One-to-one}(f) \leftrightarrow (\forall x)(\forall y)(\forall z)\Big(\big((x, y) \in f \,\&(z, y) \in f\big) \to x = z\Big). \tag{18.11}$$

It is clear that for a one-to-one function, its inverse as a relation is a function. In this case, one calls f^{-1} the inverse function of f, or simply the inverse of f.[18.4]

Is is clear from equation (18.5 that if f is one-to-one, then

$$\mathrm{dom}(f^{-1}) = \mathrm{ra}(f) \qquad \text{and} \qquad \mathrm{ra}(f^{-1}) = \mathrm{dom}(f). \tag{18.12}$$

If f is a one-to-one function from A onto B, this means that f^{-1} is a function from B to A. In particular, $\mathrm{dom}(f) = A$ and $\mathrm{dom}(f^{-1}) = B$.

18.5 Composition of a function and its inverse

For a set S, let id_S denote the identity function on S; that is

$$\mathrm{id}_S = \{(x, x) : x \in S\}. \tag{18.13}$$

We have

Theorem 18.1. *Let $f : A \to B$ be one-to-one and onto B, then*

$$f \circ f^{-1} = \mathrm{id}_B \qquad \textit{and} \qquad f^{-1} \circ f = \mathrm{id}_A\,.$$

[18.4] If f is not one-to-one and one wants to talk about its inverse, one always needs to say inverse as a relation. When one simply talks about the inverse of a function, one always assumes that f is one-to-one. Above, we did not strictly observe this requiremnt, since we wanted to explain the matter without unnecessary complications, but from now on, we will strictly observe this requirement.

Proof. Assume $g = f^{-1}$ According to the equations in formula (18.12) we have $\mathrm{dom}(f) = \mathrm{ra}(g)$ and $\mathrm{ra}(f) = \mathrm{dom}(g)$. Hence, by formula (18.8) we have $\mathrm{dom}(f \circ g) = \mathrm{dom}(g) = B$ and we have $\mathrm{dom}(g \circ f) = \mathrm{dom}(f) = A$.

Further, formulas (18.7) and (18.9) say that $(x, x) \in f \circ g$ if $x \in \mathrm{dom}(f \circ g)$; since $f \circ g$ is a function, we cannot have $(x, y) \in f \circ g$ for $y \neq x$ in this case. So $f \circ g$ is an identity function. The same argument with f and g interchanged shows that $g \circ f$ is an identity function. With the domains given above, this means that $f \circ g = \mathrm{id}_B$ and $g \circ f = \mathrm{id}_A$, as we wanted to show, □

The converse is also true:

Theorem 18.2. *Let $f : A \to B$ and $g : B \to A$ be functions such that*

$$f \circ g = \mathrm{id}_B \qquad \textit{and} \qquad g \circ f = \mathrm{id}_A \,.$$

Then f is one-to-one and onto B, and $g = f^{-1}$.

Proof. According to formula (18.9), we need to show that if $(x, y) \in f$ then $(y, x) \in g$, and, conversely if $(y, x) \in g$ then $(x, y) \in f$.

Assume first that $(x, y) \in f$. Then $x \in A$ and $y \in B$; by the latter, g is defined at y. Let $z \in A$ be such that $(y, z) \in g$. Then $(x, z) \in g \circ f$ according to (18.7). Given that $g \circ f = \mathrm{id}_A$, we have $(x, z) \in \mathrm{id}_A$. Since, as we remarked above, we have $x \in A$, we also have $(x, x) \in \mathrm{id}_A$. Since, given x, the z such that $(x, z) \in \mathrm{id}_A$ must be unique, we have $z = x$. Hence, recalling that $(y, z) \in g$, we have $(y, x) \in g$, as we wanted to show.

Assume now that $(y, x) \in g$. Then $(x, y) \in f$ can be shown in exactly the same way, with the roles of f and g interchanged.

Finally, to show that f is onto B, observe that for any $y \in B$ there is an x such that $(y, x) \in g$, since $\mathrm{dom}(g) = B$. Since we then have $(x, y) \in f$, it follows that $B \subset \mathrm{ra}(f)$. Since we have $\mathrm{ra}(f) \subset B$ by our assumptions, $B = \mathrm{ra}(f)$ follows. □

Corollary 18.1. *Let f and g be functions. If $g = f^{-1}$ then $f = g^{-1}$*

Of course, by the assumption that $g = f^{-1}$, we tacitly assume that f is one-to-one,

Proof. Let $A = \mathrm{dom}(f)$ and $B = \mathrm{ra}(f)$. By Theorem 18.1 the functions f and g satisfy the assumptions of Theorem 18.2 (cf. also (18.12). Since in the latter theorem, the roles of f and g are interchangeable, it follows by this latter theorem that $f = g^{-1}$. □

18.6 Inverse of a composition

Lemma 18.3. *Assume f and g are one-to-one functions, and assume that $(x, y) \in f \circ g$. Then $(y, x) \in g^{-1} \circ f^{-1}$.*

Note that nothing is assumed here about the domains and ranges of f and g here.

Proof. Given that $(x, y) \in f \circ g$, by (18.7), there is a z such that $(x, z) \in g$ and $(z, y) \in f$. By (18.9), for the same z we have $(y, z) \in f^{-1}$ and $(z, x) \in g^{-1}$; hence, by (18.7) again, $(y, x) \in g^{-1} \circ f^{-1}$, as we wanted to show. □

Corollary 18.2. *Assume f and g are one-to-one functions, Then $(f \circ g)^{-1} = g^{-1} \circ f^{-1}$.*

Of course, this implies that $f \circ g$ is one-to-one, but there is no need to show that separately.

Proof. Using Lemma 18.3 with f^{-1} replacing f and g^{-1} replacing g, in view of Corollary 18.1 we obtain the conclusion that if $(y,x) \in g^{-1} \circ f^{-1}$ then $(x,y) \in f \circ g$. Adding to this the original conclusion of Lemma 18.3, we can see that

$$(x,y) \in f \circ g \leftrightarrow (y,x) \in g^{-1} \circ f^{-1}$$

for every x and y. As $g^{-1} \circ f^{-1}$ is a function, it follows from formula (18.9) that $(f \circ g)^{-1} = g^{-1} \circ f^{-1}$, which is what we wanted to prove. □

18.7 Restriction

Given a function $f : A \to B$ and a set $C \subset A$, sometimes one wants to consider f only as a function on C. One can do this by defining the restriction $f \restriction C$ of f to C:

$$f \restriction C \stackrel{def}{=} f \cap (C \times B). \tag{18.14}$$

That is, $f \restriction C(x) = f(x)$ if $x \in C$, and it is undefined if $x \notin C$.

18.8 Injective, bijective, and surjective functions

These terms used in [2], and the related terms *injection*, *bijection*, and *surjection* originated in the French mathematical literature at a time when good English terms were available to describe these concepts, and it is not necessary to use them to replace the corresponding English terms. The term surjective if particularly troublesome, since it appears to take the *codomain* as being part of the function. The book [2, § 10.2, p. 257] defines surjective and onto as being synonymous; if so, then what is the point of using the term surjective? In some mathematical writing, the codomain (see Subsection 18.2) is considered a part of the description of the function; if so, then the use of the term surjective is defensible. But this viewpoint is never adopted in set theory, and we did not adopt it either. Without this, one would often need to say that "f is surjective to B," but then it is much simpler and clearer to say that "f is onto to B." Our point of view, the point of view generally accepted in set theory, also makes the distiction between injective and bijective without an explicit identification of the codomain; while the codomain is useful in certain practical discussions, from a set-theoretical point of view it is an artificial concept. For all these reasons, we strongly discourage the use of the terms injective, bijective, and surjective, except in specialized contexts.[18.5]

18.9 Reading

[2, Chapter 10, pp. 251–277].

18.10 Homework

[2, Chapter 10, pp. 251–277]. p. 255, 10.3, 10.5, 10.7, p. 258, 10.21, 10.27, p. 262, 10.31, p. 266, 10.37, 10.41, 10.43, 10.47, p. 273, 10.51, 10.55, 10.61.

[18.5] Such specialized contexts include category theory when the category of sets are discussed. In this case, the objects are sets, and the morphisms are functions. The source of a morphism is called domain, and the target is called codomain. In this case, the codomain is an important part of the discussion. Any discussion of category theory is, however, beyond the scope of these notes.

19 Cardinalities

The sets A and B are called *equinumerous* if there is a one-to-one function $f : A \to B$ onto B. Writing $\mathbb{N} = \{1, 2, 3, \ldots\}$, we say that A is *countably infinity* if $\mathbb{N}$ and A are equinumerous. A is said to be *countable* if A is finite or countably infinite.

Lemma 19.1. *If $X \subset \mathbb{N}$ is infinite then X is countably infinite.*

Proof. Define the function $f : X \to \mathbb{N}$ as follows. For each $n \in X$ put

$$f(n) = \min\{i \in \mathbb{N} : i \neq f(k) \text{ for } k < n \text{ with } k \in X\}.$$

Observe that this is a recursive definition; that is, the definition of $f(n)$ relies on the definition of $f(k)$ for $k < n$ with $k \in X$.[19.1]

We claim that f is one-to-one. In fact, if $n, k \in X$ and $k < n$, then the definition explicitly asserts that $f(n) \neq f(k)$. Further, we claim that f is onto $\mathbb{N}$. In fact, assume, on the contrary, that, for some $m \in \mathbb{N}$, there is no $k \in X$ such that $f(k) = m$. Then, for every $n \in X$ we have

$$m \in \{i \in \mathbb{N} : i \neq f(k) \text{ for } k < n \text{ with } k \in X\}.$$

As $f(n)$ is the least element of the set on the right-hand side, this implies that $f(n) \leq m$. That is, $f(n) \leq m$ for every $n \in X$. Since f is one-to-one and X is infinite, this is not possible.[19.2] □

Corollary 19.1. *Assume A is countably infinite and $B \subseteq A$ is infinite. Then B is countably infinite.*

Proof. Let $f : \mathbb{N} \to A$ be a one-to-one function onto A. Put

$$X = \{n \in \mathbb{N} : f(n) \in B\}.$$

Then X is infinite, so it is countably infinite by the above lemma. Let $g : \mathbb{N} \to X$ be a one-to-one function onto X. Then the function $f \circ g : \mathbb{N} \to A$ is[19.3] one-to-one and onto B, showing that B is countably infinite. □

Lemma 19.2. *The set $\mathbb{Z}$ of all integers is countably infinite.*

Proof. For a real number x, write $\lfloor x \rfloor$ for the largest integer $\leq x$. The function $f : \mathbb{N} \to \mathbb{Z}$ defined by

$$f(n) = (-1)^n \left\lfloor \frac{n}{2} \right\rfloor \qquad (n \in \mathbb{N})$$

is one-to-one and onto $\mathbb{Z}$. Indeed, we have $f(1) = 0$, for $k \in \mathbb{N}$ we have $f(2k) = k$ and $f(2k + 1) = -k$. □

[19.1] One might reflect that this definition gives $f(n) = 1$ for the least element of n of X, in which case the restriction on i after the colon is vacuous, since there is no $k < 1$ with $k \in X$; that is, the clause after the colon is true for every $i \in \mathbb{N}$ in this case.

[19.2] If we take m to be the least integer for which no $k \in X$ exists such that $f(k) = m$, then we can in fact conclude by this argument that the range of the function f is the set $\{1, 2, \ldots, m - 1\}$; i.e., that X has exactly $m - 1$ elements.

[19.3] We could have written $f \circ g : \mathbb{N} \to B$ instead of $f \circ g : \mathbb{N} \to A$, since it is easy to verify that all values of $f \circ g$ are in B. The emphasis here, however, is on the word "verify"; however easy the verification of this fact is, to see that the values of $f \circ g$ are in A can be seen much more directly, since all values of f are in A.

Note. We have

$$\left\lfloor \frac{n}{2} \right\rfloor = \frac{n}{2} + \frac{(-1)^n - 1}{4}.$$

Indeed, for even n, the right-hand side gives $n/2$, and for odd n it gives $(n-1)/2$. Hence we can also define the above function f as

$$f(n) = (-1)^n \frac{2n + (-1)^n - 1}{4} \qquad (n \in \mathbb{N}).$$

Lemma 19.3. *The Cartesian product $\mathbb{N} \times \mathbb{N}$ is countably infinite.*

Proof. A one-to-one function $f : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ can be defined as follows. Given $(m, n) \in \mathbb{N} \times \mathbb{N}$, put[19.4]

$$f(m, n) = \frac{(m + n - 1)(m + n)}{2} + n - 1$$

for $m, n \in \mathbb{N}$.

We claim that f is one-to-one. To show this, let (m, n) and (k, l) be two pairs of positive integers such that $(m, n) \neq (k, l)$.[19.5] Without loss of generality, we may assume that $m + n \leq k + l$. If $m + n = k + l$ then we must have $n \neq l$, so

$$\begin{aligned} f(m, n) &= \frac{(m + n - 1)(m + n)}{2} + n - 1 = \frac{(k + l - 1)(k + l)}{2} + n - 1 \\ &\neq \frac{(k + l - 1)(k + l)}{2} + l - 1 = f(k, l). \end{aligned}$$

If $m + n < k + l$ then we have $m + n \leq k + l - 1$ and $m + n + 1 \leq k + l$, and so

$$\begin{aligned} f(m, n) &= \frac{(m + n - 1)(m + n)}{2} + n - 1 < \frac{(m + n - 1)(m + n)}{2} + m + n - 1 \\ &= \frac{(m + n - 1)(m + n) + 2(m + n)}{2} - 1 = \frac{(m + n + 1)(m + n)}{2} - 1 \\ &= \frac{(m + n)(m + n + 1)}{2} - 1 \leq \frac{(k + l - 1)(k + l)}{2} - 1 \\ &< \frac{(k + l - 1)(k + l)}{2} + l - 1 = f(k, l), \end{aligned}$$

so again $f(m, n) \neq f(k, l)$. This shows that f is one-to-one.

Write

$$X = \{n \in \mathbb{N} : n = f(k, l) \text{ for some } k, l \in \mathbb{N}\}.$$

Then $X \subseteq \mathbb{N}$ is infinite, and so X is countably infinite. Since $f : \mathbb{N} \times \mathbb{N} \to X$ is a one-to-one mapping that is onto X, this shows that $\mathbb{N} \times \mathbb{N}$ is also countably infinite, which is what we wanted to prove. □

Note. The function f defined in the last proof is not onto $\mathbb{N}$. The function $g : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ defined by

$$g(m, n) = \frac{(m + n - 2)(m + n - 1)}{2} + n$$

for $m, n \in \mathbb{N}$ is onto $\mathbb{N}$. The calculations showing that g is one-to-one are slightly more complicated than the ones showing that f is one-to-one, and showing that g is onto $\mathbb{N}$ requires extra effort. Since it was not important in the above proof that f be onto $\mathbb{N}$, it was simpler to work with the function f instead of g.

[19.4] Strictly speaking, an element of $\mathbb{N} \times \mathbb{N}$ has the form (m, n) for positive integers m and n. The value of f at such an element should be denoted as $f\big((m, n)\big)$. However, it visually more pleasing to use the notation $f(m, n)$ at the price of some inaccuracy.

[19.5] That is, $m \neq n$ or $k \neq l$.

Corollary 19.2. *If A and B are countably infinite sets then $A \times B$ is also countably infinite.*

Proof. Let $f : \mathbb{N} \to A$ onto A, $g : \mathbb{N} \to B$ onto B, and $h : \mathbb{N} \to \mathbb{N} \times \mathbb{N}$ onto $\mathbb{N} \times \mathbb{N}$ be one-to-one functions. Such functions exist since A and B are countably infinite by assumption, and $\mathbb{N} \times \mathbb{N}$ is countably infinite by the last lemma.

We define the function $\phi : \mathbb{N} \to A \times B$ as follows. Given $n \in \mathbb{N}$, let $k, l \in \mathbb{N}$ be such that $h(n) = (k, l)$, and let $\phi(n) = (f(k), g(l))$. It is easy to show that ϕ is one-to-one and onto $A \times B$. □

In what follows, $\mathbb{Q}$ will denote the set of rational numbers, and $\mathbb{Q}^+$ will denote the set of positive rationals. We have

Lemma 19.4. *$\mathbb{Q}^+$ is countably infinite.*

Proof. Write

$$S = \{(m, n) : m, n \in \mathbb{N} \text{ and the greatest common divisor of } m \text{ and } n \text{ is } 1\}.$$

The set S is an infinite subset of the countably infinite set $\mathbb{N} \times \mathbb{N}$, so it is countably infinite by Corollary 19.1. The function f defined as

$$f(m, n) = \frac{m}{n} \qquad \text{for} \quad (m, n) \in S$$

is a one-to-one function from S onto $\mathbb{Q}^+$, showing that $\mathbb{Q}^+$ is also countably infinite. □

For a set A, the power set $\mathcal{P}(A)$ of A is defined as the set of all subsets of A. The following theorem and its proof is valid for any set A, be A finite or infinite; the proof is valid even when A is the empty set.[19.6] But only the case when A is infinite is of real interest, since for finite A a much more precise statement can be made.

Theorem 19.1. *Let A be an arbitrary set. Then A and $\mathcal{P}(A)$ are not equinumerous.*

Proof. We will show that there is no one-to-one function from A onto $\mathcal{P}(A)$; in fact, no function from A onto $\mathcal{P}(A)$ exist, whether or not we require it to be one-to-one. To see this, let $f : A \to \mathcal{P}(A)$ be an arbitrary function, and consider the subset C of A defined as

$$C = \{x \in A : x \notin f(x)\}.$$

Then there is no $y \in A$ for which $f(y) = C$. Indeed, for an arbitrary $y \in A$, if $y \in f(y)$ then $y \notin C$ by the definition of C, and if $y \notin f(y)$ then $y \in C$. This shows that $f(y)$ and C do not have the same elements (y is an element of exactly one of these two sets), so $f(y) \neq C$, as claimed. □

A related argument can be given that $\mathbb{N}$ and the set of real numbers, $\mathbb{R}$, are not equinumerous. On the other hand, we have the following

Theorem 19.2. *The set $\mathbb{R}$ and the interval $(-1, 1)$ are equinumerous.*

Proof. The function $f : (-1, 1) \to \mathbb{R}$ such that

$$f(x) = \frac{x}{x^2 - 1} \qquad \text{for} \quad x \in (-1, 1)$$

[19.6] As one might expect, the case when A is the empty set involves a number of vacuously true statements.

is one-to-one and onto $\mathbb{R}$. Indeed, let $y \in \mathbb{R}$ be arbitrary. We need to show that there is exactly one $x \in (-1,1)$ such that $f(x) = y$. If $y = 0$ then we have $f(x) = y$ only for $x = 0$. Assume now that $y \neq 0$. Then the equation $f(x) = y$ can be equivalently written as $y(x^2 - 1) = x$.

Observe that this latter equation makes sense for $x = \pm 1$ while the equation $f(x) = y$ does not. The important point, however, is that $x = \pm 1$ does not satisfy the latter equation, since for this choice of x the left-hand side is 0 while the right-hand side is ± 1. That is, the exceptional case of $x = \pm 1$ does not affect the equivalence of the two equations.

Keeping in mind that we assumed that $y \neq 0$, the latter equation can also be written as

$$x^2 - \frac{1}{y}x - 1 = 0.$$

This is a quadratic equation for x. We can solve this equation for x as

$$x = \frac{\frac{1}{y} \pm \sqrt{\frac{1}{y^2} + 4}}{2}.$$

Given that the discriminant (the expression under the square root) of this equation is positive, this equation has two distinct real solutions; call them x_1 and x_2. The product of these two solutions is the constant term of the equation; that is, $x_1 x_2 = -1$. Therefore $|x_1||x_2| = 1$. Given that $|x_1|, |x_2| \neq 1$, as we remarked above, one of x_1 and x_2 must be inside the interval $(-1,1)$ and the other one must be outside. That is, there is exactly one $x \in (-1,1)$ for which $f(x) = y$, as we wanted to show. □

19.1 The Cantor-Schröder-Bernstein Theorem

Consider the sets $A = \mathbb{R}$ and $B = [-1,1]$, the closed interval from -1 to 1. There there are functions $f : A \to B$ into B and $g : B \to A$ into A that are one-to-one. Namely, A is equinumerous to $(-1,1)$ according to the last theorem; so we can take f to be the one-to-one function from A onto $(-1,1)$. For the function g we can simply take the identity function on B. Since $B \subset A$, g is into A. The next theorem asserts that under these conditions A and B are equinumerous.

Theorem 19.3 (Cantor-Schröder-Bernstein Theorem)**.** *Let A and B be sets and assume there are one-to-one functions $f : A \to B$ and $g : B \to A$. Then A and B are equinumerous.*

In the theorem, it is of course not required that f be onto B or g be onto A; in fact, there would be nothing to prove if this were the case. The result is intuitively obvious if A or B are finite sets, but for infinite sets the result is not at all obvious, and a proof is needed. The proof, however, works regardless whether A or B are finite or infinite. We will give two proofs. The first proof gives a much better insight why the result is true, especially if one follows along the first proof by drawing a picture. The second proof can be presented more concisely, but it gives little insight why the result is true. The second proof is the one that is usually given in textbooks. Before giving the formal version of the first proof, we include an intuitive description.

Imagine the sets A and B as two vertical lines, A on the left, B on the right. From each point, left or right, draw one, and if possible, two edges, one forward edge to the image of the point under the function f or g (whichever is applicable), and one backward edge, to the the inverse image. The forward image always exists, the backward image may not exist, since the inverses need not be defined everywhere. These edges can be continued to complete paths containing the given point. Two different paths may not have points in common, since the functions f and g are one-to-one. A path may have a starting point, where the inverse image does not exist, or may go back indefinitely

(when it may loop back on itself, forming a cycle, or may not; a cycle will always contain an even number of edges). Select alternate edges of these paths, making sure that the starting point of the path is incident to an edge. (When the path has no starting point, it is immaterial how the edges are selected, but below we select the edge going backward from left to right, since this slightly simplifies the formal description). The selected edges form a one-to-one mapping from A to B (when redirected from left to right if necessary). The mapping is defined on all of A and is onto B for the same reason: each point is incident to a path.

We will now give a formal description of the proof.

First Proof. For a function ϕ one usually denotes by $\phi(x)$ the value of ϕ at x, but one sometimes uses the notation ϕx instead. It will be advantageous for us to use this latter notation in what follows in order to avoid having to write too many parentheses.[19.7] Denote by f^{-1} and g^{-1} the inverses of the functions f and g. For every $a \in A$, form the sequence

$$a, \quad g^{-1}a, \quad f^{-1}g^{-1}a, \quad g^{-1}f^{-1}g^{-1}a, \quad f^{-1}g^{-1}f^{-1}g^{-1}a, \ \ldots.$$

Note that g^{-1} maps a subset of A into B, and f^{-1} maps a subset of B into A. Hence $g^{-1}a$ may or may not be defined. Even if $g^{-1}a \in B$ is defined $f^{-1}g^{-1}a$ may not be defined, and if $f^{-1}g^{-1}a \in A$ is defined, the element $g^{-1}f^{-1}g^{-1}a$ may then not be defined. That is, the above sequence may terminate at some point. Call the sequence associated with an $a \in A$ in this way $\sigma(a)$.

The elements of this sequence need not be distinct; for example, we can have $f^{-1}g^{-1}a = a$, in which case the sequence never terminates, but it has only two distinct elements, a and $g^{-1}a$. When talking about the *number of elements* of this sequence, we will mean its length, and not the number of its distinct elements. For example, in case $f^{-1}g^{-1}a = a$ we will say that the above sequence has infinitely many elements, even though the number of its distinct elements is only two.

Now, define a function $h : A \to B$ as follows. For an arbitrary $a \in A$, consider the above sequence. If (i) the sequence $\sigma(a)$ has infinitely many elements, or a finite even number of elements, then put $ha = g^{-1}a$, and if (ii) the $\sigma(a)$ has an odd number of elements then put $ha = fa$. First, observe that this defines ha for every $a \in A$. Indeed, in case (i), the sequence $\sigma(a)$ has at least two elements, so $g^{-1}a$ is defined; so ha is defined in case (i); since fa is defined for every $a \in A$, ha is also defined in case (ii).

We show that h is one-to-one. To this end, let $a_1, a_2 \in A$ such that $a_1 \neq a_2$. If both ha_1 and ha_2 are defined according to clause (i), then $ha_1 = g^{-1}a_1 \neq g^{-1}a_2 = ha_2$. Similarly, if both ha_1 and ha_2 are defined according to clause (ii), then $ha_1 = fa_1 \neq fa_2 = ha_2$. So assume one of ha_1 and ha_2 is defined according to clause (i), and the other according to clause (ii). Without loss of generality, we may assume that ha_1 is defined according to clause (i) and ha_2 is defined according to clause (ii). Assume, that $ha_1 = ha_2$, i.e., $g^{-1}a_1 = fa_2$. Then $a_2 = f^{-1}g^{-1}a_1$. Hence the sequence $\sigma(a_1)$ can be written as

$$a_1, \quad g^{-1}a_1, \quad \sigma_1(a_2), \quad \sigma_2(a_2), \quad \sigma_3(a_2), \ \ldots,$$

where $\sigma_1(a_2)$, $\sigma_2(a_2)$, $\sigma_3(a_2)$, ... denote the first, second, third, ... elements of the sequence $\sigma(a_2)$. Now, $\sigma(a_2)$ has an odd number of elements, since clause (ii) was used to define ha_2, and $\sigma(a_1)$ has either an infinite number of elements or a finite even number of elements, since clause (i) was used to define ha_1. This is a contradiction, since we just saw that $\sigma(a_1)$ has exactly two more elements than $\sigma(a_2)$. This contradiction shows that h is one-to-one.

[19.7] The notation ϕx could be confusing where *juxtaposition* (i.e., placing next to each other) of letters can indicate multiplication; this is why one usually uses the notation $\phi(x)$ instead. In the present case, multiplication is not used, so there is no such danger.

To show that h is onto B, let $b \in B$ be arbitrary, and define the sequence

$$b, \quad f^{-1}b, \quad g^{-1}f^{-1}b, \quad f^{-1}g^{-1}f^{-1}b, \quad g^{-1}f^{-1}g^{-1}f^{-1}b, \ \ldots.$$

Call this sequence $\rho(b)$. Let $a_1 = gb$. Then $b = g^{-1}a_1$, and so the sequence $\sigma(a_1)$ can be written as

$$a_1, \quad \rho_1(b), \quad \rho_2(b), \quad \rho_3(b), \ \ldots,$$

where $\rho_1(b)$, $\rho_2(b)$, $\rho_3(b)$, ... denote the first, second, third, ... elements of the sequence $\rho(b)$. That is, $\sigma(a_1)$ has one more elements than $\rho(b)$. Hence, if $\rho(b)$ has an infinite number of elements or a finite odd number of elements then $\sigma(a_1)$ has either an infinite number of elements or a finite even number of elements, and so $ha_1 = g^{-1}a_1 = b$.

Assume now that $\rho(b)$ has a finite even number of elements. Then it has at least two elements, so $a_2 = f^{-1}b$ is defined. Then the elements of the sequence $\rho(b)$ can be written as

$$b, \quad \sigma_1(a_2), \quad \sigma_2(a_2), \quad \sigma_3(a_2), \ \ldots,$$

showing that $\sigma(a_2)$ has one fewer element than $\rho(b)$. That is, $\sigma(a_2)$ has an odd number of elements, and so $ha_2 = fa_2 = b$. This shows that h is onto B. □

The second proof defines the same one-to-one function $h : A \to B$, but it describes this function in a different way. After carefully reading both proofs, one should realize that the two proofs are essentially the same, presented differently.

Second Proof. Write $A_0 = A$, $B_0 = B$, and if A_n and B_n for $n \geq 0$ have been defined, put

$$A_{n+1} = \{g(b) : b \in B_n\} \quad \text{and} \quad B_{n+1} = \{f(a) : a \in A_n\}.$$

Note that since f is one-to-one, this means that, for every $n \geq 0$, $a \in A_n$ if and only if $f(a) \in B_{n+1}$.[19.8] Similarly, $b \in B_n$ is and only if $g(b) \in A_{n+1}$.

We clearly have $A_1 \subseteq A_0$ and $B_1 \subseteq B_0$. By induction, it is then easy to prove that $A_{n+1} \subseteq A_n$ and $B_{n+1} \subseteq B_n$. Indeed, if we assume that for some $n > 1$ we have both $A_n \subseteq A_{n-1}$ and $B_n \subseteq B_{n-1}$, then $A_{n+1} \subseteq A_n$ follows from the latter of these relations, and $B_{n+1} \subseteq B_n$ follows from the former. Let

$$C = \bigcap_{n=0}^{\infty} A_n.$$

Define the function h on A as follows. If (i) $a \in A_n \setminus A_{n+1}$ for some odd n or $a \in C$ then put $h(a) = g^{-1}(a)$, where g^{-1} is the inverse of g, and if (ii) $a \in A_n \setminus A_{n+1}$ for some even n, then put $h(a) = f(a)$. We then have to show that (1) the definition of h is meaningful, (2) h is defined for every $a \in A$, (3) h is one-to-one, and (4) h is onto B.

As for (1) and (2), let $a \in A$. Then either $a \in C$ or there is an k for which $a \notin A_k$. Assume that $a \notin C$, and let $k \geq 0$ be the least integer for which $a \notin A_k$. Then $a \in A_m$ for every nonnegative integer $m < k$ and $a \notin A_m$ for every $m \geq k$. Hence for the integer n such that $a \in A_n \setminus A_{n+1}$ we must have $n = k-1$. That is, this integer n is uniquely determined. This makes the definitions given in (i) and (ii) meaningful, except that we need to show that $g^{-1}(a)$ is defined in case (i). However, $g^{-1}(a)$ is defined unless $a \in A_0 \setminus A_1$, and $a \in A_0 \setminus A_1$ is not true in case (i).

[19.8] The important point here is that if $a \notin A_n$ then $f(a) \notin B_{n+1}$. Indeed, if we had $f(a) \in B_{n+1}$, there would have to be an $a' \in A_n$ for which $f(a') = f(a)$. But this is not possible, since f is one-to-one, and so we would have to have $a' = a$, but $a \notin A_n$.

As for (3), assume $h(a_1) = h(a_2)$ for some $a_1, a_2 \in A$ such that $a_1 \neq a_2$. Since f and g^{-1} are one-to-one, this is not possible if both $h(a_1)$ and $h(a_2)$ are both defined by clause (i) or if they are both defined by clause (ii). Assume, therefore, that they are defined by different clauses. Without loss of generality, we may assume that $h(a_1)$ is defined by clause (i) and $h(a_2)$ is defined by clause (ii). We then have $g^{-1}(a_1) = h(a_1) = h(a_2) = f(a_2)$. Since we used clause (ii) to define $h(a_2)$, we have $a_2 \in A_n \setminus A_{n+1}$ for some even n. Writing $b = f(a_2)$, we have $b \in B_{n+1} \setminus B_{n+2}$. Since we also have $b = g^{-1}(a_1)$, i.e., $a_1 = g(b)$, it follows that $a_1 \in A_{n+2} \setminus A_{n+3}$. As $n+2$ is even, this contradicts the assumption that $h(a_1)$ was defined according to clause (i).

As to (4), let $b \in B$. If $b \in B_n$ for all $n \geq 0$ then, writing $a = g(b)$, we have $a \in A_n$ for all $n \geq 0$, and so $a \in C$. Therefore, $h(a) = g^{-1}(a) = b$ according to clause (i). Assume that this is not the case, and let $n \geq 0$ be the unique integer such that $b \in B_n \setminus B_{n+1}$. If n is even then, with $a = g(b)$, we have $a \in A_{n+1} \setminus A_{n+2}$. Since $n+1$ is odd, we have $h(a) = g^{-1}(a) = b$ according to clause (i). If n is odd, then $n \geq 1$; thus $b \in B_n \subseteq B_1$, which implies that $b = f(a)$ for some $a \in A$. Then $a \in A_{n-1} \setminus A_n$, and $h(a) = f(a) = b$ according to clause (ii). This completes the proof. □

19.2 Reading

[2, Chapter 11, pp. 278–302].

19.3 Homework

[2, Chapter 11, pp. 278–302]. p. 280, 11.1, p. 258, 11.9, 11,15, 11.18 (even numbered, no solution in textbook), 11.19, p. 292, 11.23, p. 295, 11.29, 11.31, p. 300, 11.33, 11.35.

20 Paradoxes in mathematics

20.1 Cantor's paradox

When Georg Cantor found this result, there was no restriction on what he or his followers considered a set (there were people like Leopold Kronecker who did not find Cantor's arguments acceptable. The world soon soured on Cantor's discovery. In fact, it was Cantor himself that discovered a contradiction in around 1899. Writing V for the set of all sets, we cannot have

$$|\mathcal{P}(V)| > |V|.$$

since every element of $\mathcal{P}(V)$ is already an element of V. This statement, showing that Theorem 19.1, even though proved apparently rigorously, cannot be valid for V replacing A, is called Cantor's paradox; a paradox is a self-contradictory statement.

The set theory invented by Cantor is often called naive set theory (though I prefer the name intuitive set theory, i.e., a set theory based on intuition, rather than on a rigorous foundation). The discovery of paradoxes in set theory was very disconcerting to mathematician, since set theory became an important part of mathematics. One of the greatest mathematicians at the time, David Hilbert wrote that "Aus dem Paradies, das Cantor uns geschaffen, soll uns niemand vertreiben können" (From the paradise that Cantor created for us, no-one shall be able to chase us out).

20.2 Real classes

The resolution of this paradox was given by aximatic set theory, in which large collections are called (real) classes, and small collections are sets; we will say more about axiomatic set theory later. Here

large and small are relative term, so we will give a more precise explanation: all collections are considered classes, but some collections will be considered sets. Classes (and sets) can only have sets as elements. Real classes (classes that are not sets) cannot be elements of classes. This will avoid Cantor's paradox, but it is not at all clear whether paradoxes will re-appear in different forms. '

20.3 Russell's paradox

One of the most famous paradoxes is due to Bertrand Russell. It starts out with saying that for a set x, we may or may not have $x \in x$.[20.1] Now, consider the set R defined as

$$R = \{x : x \notin x\},$$

where x runs over all sets. That is, R is the set of all sets x for which $x \notin x$. Then one asks the question whether or not $R \in R$. This question has no answer, since if the answer is yes, that is, if $R \in R$, then the definition of R with $x = R$ says that $R \notin R$ and if the answer is no, that is, if $R \notin R$, then the definition of R with $x = R$ says that $R \in R$. Russel was so upset with his own paradox that he decided to write a monumental three-volume work with Alfred North Whitehead[20.2] in which, as detractors of the work like to say, it took 1500 pages to prove that $1 + 1 = 2$. Their aim was to put mathematics on a solid foundation; but, as Kurt Gödel showed in 1931, the goal they tried to accomplish is not attainable.

20.4 Epimenides paradox

Logical paradoxes have been around for a long time. Epimenides was a Cretan (that is, from the island of Crete in the Mediterranean sea) in the 7th or 6th century BC (perhaps he was a mythical rather than a real figure) who said that "all Cretans are liars," therefore he himself is a liar. But that means the statement is not true, so he is not a liar after all. But then the statement is true, so he is in fact a liar. So which is it? Even though this is not formulated in terms of set theory, the analogy with Russell's paradox is unmistakable. A somewhat similar paradox motivated by Russell's paradox is the barber's paradox, described as saying "the barber is the man who shaves those that do not shave themselves."[20.3] The question is, does the barber shave himself?

20.5 Berry's paradox

The paradox of G.G. Berry, a junior librarian at Oxford describes an integer as "The smallest positive integer not definable in under sixty letters" (this phrase conain 57 letters). So, which integer it defines? The problem is that if it in fact defines an integer, then the phrase says that that integer cannot be defined in fewer than sixty letters. Obviously, the sentence itself is contradictory. The resolution of this paradox is that mathematical statements must be described in a formal mathematical language; natural languages such as English are not sufficient to make clear statements. Such a formal language has in fact been developed. For example, the statement that an integer p is prime if it greater than 1 and it cannot be represented as a product of two integers u and v with $1 < u < p$ and $1 < v < p$ is described by the formal statement

$$\text{prime}(p) \leftrightarrow \forall u \forall v \big((1 < u < p \,\&\, 1 < v < p) \rightarrow p \neq uv \big),$$

[20.1] An axiom called the Axiom of Foundation, invented by John von Neumann, explicitly disallows the possibility that $x \in x$ for any set x.
[20.2] A building is named after him at Brooklyn College
[20.3] Here shaving means shaving facial hair, so the statement is traditionally only about men – though nowadays one seems to treading dangerous grounds for making such statements.

where the variables p, u, and v run over integers. The literal translation of this fomral statement is the following: p is a prime (expressed formally as prime(p)) if and only if (expressed as $\leftrightarrow$) for all u (expressed formally as $\forall u$) and for all v (expressed as $\forall v$)[20.4]), if $1 < u < p$ and (expressed as &) $1 < v < p$ then (the "if ... then" is expressed as $\rightarrow$) $p \neq uv$). Note that with the placements of parentheses in the formula one can express more clearly what is meant than is possible in a natural language without sometimes complicated circumlocutions.

20.6 Axiomatic set theory

It was strongly felt by some mathematicians that in spite of the paradoxes described above, set theory should not be discarded, especially since the paradoxical statements were different from the mathematically useful statements, but nobody really knew where the boundary between the two lied. We mentioned the attempt by Russell and Whitehead to resolve the problem; a humorous description of this history can be found in the essay The Greatest Math mistake of the Century.[20.5] The theory created by Russell and Whitehead was awkward to use in practical mathematics. An axiom system for set theory of practical use was created by Ernst Zermelo. The theory created by Russell and Whitehead was awkward to use in practical mathematics. An axiom system for set theory of practical use was created by Ernst Zermelo. In axiomatic set theory, the only kinds of things are sets; that is, the elemets of a set are sets themselves (in class, we discussed how to define the integers as sets of earlier integers.[20.6]

We will briefly discuss some of Zermelo's axioms. The *axiom of extensionality* says that two sets that sets are equal if and only if thet have the same elements (this was mentioned in class). The *axiom of power set* states the existence of the set of all subsets (power set) of a set. The *axiom of infinity* asserts the existence of a specific infinite set (and not just infinite sets in general). A remarkable axiom is the *axiom of choice*, basically saying that we can pick elements from an infinite number of sets. More precisely, it asserts the following: given a set x. there is a function f such that for each nonempty $y \in x$ we have $f(y) \in y$.[20.7] Formally,

$$(\forall x)\big(\exists f : \text{Function}(f)\big)\big(\forall y \in x\big)\Big(\big(y \neq \emptyset \rightarrow (\exists z \in y)\big((y,z) \in f\big)\big)\Big); \tag{20.1}$$

see formula (18.1) for the definition of Function(f). Interestingly, this axiom caused a lot of controversy, even though on first hearing it seems completely natural. However, its consequences are so striking that many mathematicians started to doubt its validity (or its safety, in that they questioned whether it can cause contradictions, as earlier considerations in naive set theory caused contradictions). It turns out that it is completely safe, as Gödel showed in 1938 (to be discussed in more detail below).

[20.4]Note that this no separate symbol is used to express the word "and." This is because the word "and" here is not used in the logic sense, as in "the sun shines *and* it is warm," but in the sense of enumeration as in "ham *and* eggs."

[20.5]The html file that contains this essay has the following copyright notice (commented out in the html, so it is not visible for the reader without looking at the source file): this website is copyrighted by Paul Cox, all rights reserved. The use of this material for commercial purposes without permission, including the posting of advertisements on this site, is a violation of that copyright and may result in civil prosecution. Unfortunately, the last time I found this file on the internet was in 2012, and it has disappeared before; so I felt I need to download it, so it will not be lost. I am posting the website here for non-commercial use, since I feel it has great educational value, and without my posting it it would disappear from public view. See the site for more.

[20.6]In set theory, by *integers* one means nonnegative integers. As we mentioned, $0 = \emptyset$, $1 = \{0\} = \{\emptyset\}$, $2 = \{0,1\} = \{\emptyset, \{0\}\}$, $3 = \{0,1,2\} = \{\emptyset, \{0\}, \{\emptyset, \{0\}\}\}$, ...; these are called the *von Neumann ordinals*; ordinals, and not integers, because the construction continues after all integers have been described this way.

[20.7]The intended domain of f is the set of nonempty elements of x, but this need not be stated explicitly. In set theory, as discussed in Section 18, functions themselves are considered sets, namely sets of pairs. For example, the function $f(x) = x^2$ on the set $\mathbb{R}$ of all real numbers is considered the set $f = \{(x, x^2) : x \in \mathbb{R}\}$.

20.7 The axiom of replacement

Zermelo's axiom system had one serious deficiency in that it could not prove that the collection containing the elements $\mathbb{N} = \{0, 1, 2, 3, \ldots\}$, $\mathcal{P}(\mathbb{N})$, $\mathcal{P}(\mathcal{P}(\mathbb{N}))$, $\mathcal{P}(\mathcal{P}(\mathcal{P}(\mathbb{N})))$, $\mathcal{P}(\mathcal{P}(\mathcal{P}(\mathcal{P}(\mathbb{N}))))$, ..., ad infinitum,[20.8] is a set; Cantor certainly considered this collection a set.

This definiency was cured by Abraham (Adolf) Fraenkel in 1922 by adding what is called the *axiom replacement*, which, formulated somewhat loosely, assers that if $\phi(x, y)$ is a formula with two distinguished variables x and y[20.9] such that, given a set z, for each $x \in z$ there is exactly one y with $\Phi(x, y)$ then the collection

$$\{y : x \in z \quad \text{and} \quad \phi(x, y)\} \tag{20.2}$$

is a set.

In the formal framework, the axiom of replacement is not a single axiom; it is a group of infinitely many axioms, called an axiom scheme, one for each formula $\phi(x, y)$. The axiom scheme of replacement greatly increases the power of Zermelo's set theory. Zermelo's set theory with the axiom of choice and with the axiom of choice and with the axiom scheme of replacement is usually called ZFC (Z for Zermelo, F for Fraenkel, and C for Choice); it is the most frequently used axiom system for set theory.

20.8 Other axiom systems of set theory

Initially, the fact that ZFC had infinitely many axioms was considered a disadvantage. In 1925, John von Neumann introduced classes for large collection. In this version of set theory, all collections are classes, and those that are allowed to be elements of other classes are called sets. With the aid of this framework, he was able to reformulate set theory with finitely many axioms. Soon afterward, Paul Bernays and Kurt Gödel reformulated von Neumann's system that is now called the von Neumann–Bernays–Gödel set theory (also called Gödel–Bernays set theory).

20.9 Hilbert's program and incompleteness

As a reaction to the contradictions in naive set theory, David Hilbert proposed a solution to this crisis by grounding mathematics on a finite set of axioms which could be proved *consistent*, i.e., free of contradictions, and also complete in the sense that true mathematical statement can be proved in this system.

In 1931, Kurt Gödel surprised the mathemacal world by proving that Hilbert's program cannot be accomplished in that he showed that given any axiom system that is strong enough to contain the traditional system of axioms, due to Giuseppe Peano, is incomplete in the sense that there are true statements about integers that cannot be proved about integers; and, worse yet, Peano's system cannot be proved to be free of contradictions inside Peano's system.

The proof of consistency, i.e., of being free from contradiction, needs to be accomplished inside Peano's system, or in a weaker system that is known to be consistent. Gödel proved his result under the assumption of consistency of Peano's system. This is important, since Peano's system being inconsistent (containing a contradiction), i.e., if a false statement such as $0 = 1$ can be proved in'the system, then the system cannot be relied on (in fact, if a false statement can be proved in the system, then everything, true of false, can be proved in the system).[20.10]

[20.8] Latin for "to infinity," meaning that the list is infinitely long.

[20.9] It may have other variables that assume fixed values.

[20.10] In actual fact, Gödel used a somewhat stronger assumption than the consistency of Peano arithmetic; a few years

20.10 The continuum hypothesis

Cantor at the time of creating his set theory, asked the question whether there are infinite sets that have cardinality greater than that of $\mathbb{N}$, yet smaller than $\mathcal{P}(\mathbb{N})$.

The question whether cardinalities are comparable is decided affirmatively under the axiom of choice. That is, given any two sets A and B, there is a one-to-one mapping $f : A \to B$ (in which case, $|A| \geq |B|$) or there is a a one-to-one mapping $g : B \to A$ (in which case, $|B| \geq |A|$). Without the axiom of choice, one can prove that if there is a one-to-one mapping $f : A \to B$ and also there is a one-to-one mapping $g : B \to A$, then the is a one-to-one function $h : A \to B$ that is onto B. (That is, if $|B| \leq |A|$ and $|A| \leq |B|$ then $|A| = |B|$.

In 1938, Gödel proved that if there are no contradictions in ZF (Zermelo–Fraenkel set theory without the axiom of choice) then there are no contradictions in ZFC (Zermelo–Fraenkel set theory with the axiom choice) even with the continuum hypohesis added. This is what we meant by saying that the axiom of choice is a harmless assumption.

20.11 Independence of the continuum hypothesis

In 1964, Paul J. Cohen proved that the continuum hypothesis is independent of ZFC; that is (assuming ZF is consistent) one cannot prove the continuum hypothesis in ZFC. This means that ZFC set theory cannot decide the continuum hypothesis: it cannot prove that the continuum hypothesis is true, and it cannot prove that it is false. Given that the continuum hypothesis has very interesting consequences, this is a disconcerting situation.

The von Neumann–Bernays–Gödel (NBG) set theory cannot decide either whether the continuum hypothesis is true or false; this is no surprise, since in a technical sense NGB and ZFC are equivalent.[20.11]

20.12 Computers

In 1928, David Hilbert and Wilhelm Ackermann formulated the Entsheidungsproblem[20.12] in that they asked whether one can design a systematic method of calculation (algorithm) that can decide whether a mathematical statement is true.[20.13] In 1936, Alonzo Church and Alan Turing solved negatively, by showing that there is no such method. They gave radically different solutions: Church created λ-calculus, on which the LISP programming language is based. LISP played a very important role in the development of artificial intelligence.

Turing created a theoretical machine called the Turing machine that was very important in the development of computers. Turing showed that it is possible to create a universal machine that can do any calculation doable on other Turing machine. This led to the idea that it is in fact possible to build an actual universal machine machine that can do all calculations that can be done at all (except for resource constraints, such as time and memory; Turing's theoretical machine did not have such constraints).

---

later, John Barkley Rosser showed that this stronger assumption can be replaced with the assumption of consistency only.

[20.11] NBG is a conservative extension of ZFC. That is, any statement that can be formulated in the language of ZFC and provable in NBG is also provable in ZFC. Note that the language of NBG is richer in that it can formulate statements that cannot be formulated in the language of ZFC.

[20.12] German for decision problem; the German word is used even in the English language literature.

[20.13] The statement they used is a formal mathematical language similar to the one we used to define prime numbers above. Instead of true, they asked if the statement is universally valid, that is, whether it is true under any interpretation of the symbols in it – we want to avoid a rigorous explanation for the sake of simplicity.

Turing designed a theoretical machine called Turing machine, which played a great role in the development of actual computers,

Building on Turing's ideas, John von Neumann worked out the design principles of the a stored program computer, and such a computer was built under his direction from 1945 to 1951 at the Institute for Advanced Study in Princeton, New Jersey. The first electronic compter, the ENIAC (Electronic Numerical Integrator and Computer) was built built earlier at the University of Pennsylvania; it was completed in late 1945. The ENIAC was not a stored program computer; it was programmed with plugboards, and it took about eight hours to set up a plugboard for a new computation. In 1948, under John von Neumann's guidance, modificaions were made to the ENIAC to make it function as a stored program computer. The design of nearly all computers today are based on von Neumann's ideas, and von Neumann architecture refers to the conceptual model of this design.

21 The Axiom of Completeness of the real numbers

A *cut* is a pair (A, B) such that A and B are nonempty subsets of the set $\mathbb{R}$ of real numbers with $A \cup B = \mathbb{R}$, and such that for every $x \in A$ and every $y \in B$ we have $x < y$. If (A, B) is a cut, it is clear that the sets A and B must be disjoint (since every element of A is less than every element of B). For a cut (A, B) and a real number t we say that the cut *determines* the number t if for every $x \in A$ we have $x \leq t$ and for every $y \in B$ we have $t \leq y$.

For example, a cut that determines the number 2 is the pair (A, B) with $A = \{t : t \leq 2\}$ and $B = \{t : t > 2\}$. Another cut that determines the number 2 is the pair (C, D) with $C = \{t : t < 2\}$ and $D = \{t : t \geq 2\}$.

It is clear that a cut cannot be determine more than one number. Assume, on the contrary, that the cut (A, B) determines the numbers t_1 and t_2 and $t_1 \neq t_2$. Without loss of generality, we may assume that $t_1 < t_2$. Then the number $c = (t_1 + t_2)/2$ cannot belong to A, since we cannot have $t_1 < x$ for any element of A, whereas $t_1 < c$. Similarly $c = (t_1 + t_2)/2$ cannot be an element of B, since we cannot have $y < t_2$ for any element of B, whereas $c < t_2$.

Furthermore, the above example is of a general character, in that for every real number u there are exactly two cuts, (A, B) and (C, D) that determine u, where $A = \{t : t \leq u\}$ and $B = \{t : t > u\}$. $C = \{t : t < u\}$ and $D = \{t : t \geq u\}$. The Axiom of Completeness is an important property of real numbers:

Axiom 21.1 (Axiom of Completeness)**.** Every cut determines a real number.

Ordinarily, one does not expect to prove this statement, since axioms are basic statements that one does not prove. However, one can prove the Axiom of Completeness if one defines the real numbers as infinite decimals. [21.1]

Proof of the Axiom of Completeness. Let (A, B) be a cut. Assume first that A contains a positive number. Let n be a positive integer, and let $x^{(n)} = x_0^{(n)}.x_1^{(n)}x_2^{(n)} \ldots x_n^{(n)}$ be the largest decimal fraction with n digits after the decimal point such that $x^{(n)}$ is less than or equal to every element of B; here $x_0^{(n)}$ is a nonnegative integer, and, for each i with $1 \leq i \leq n$, $x_i^{(n)}$ is a digit, i.e., one of the

[21.1] However, defining the reals numbers as infinite decimals is inelegant in that the decimal number system came about only by a historical accident. One could also define the reals, however, as infinite binary fractions; i.e., "decimal" fractions in the binary number system. The binary number system is somewhat more natural than the decimal system in that it is the simplest of all number systems.

numbers 0, 1, ..., 9. [21.2] There is such an $x^{(n)}$, since the set of decimal fractions with n digits after the decimal point that are nonnegative and less than or equal to every element of B is finite. [21.3]

It is easy to see that if m and n are positive integers and $m < n$ then for each i with $0 \leq i \leq m$ we have $x_i^{(m)} = x_i^{(n)}$. Indeed, we cannot have $x_i^{(m)} < x_i^{(n)}$. for any i with $0 \leq i \leq m$ since then the number $x' = x_0^{(n)}.x_1^{(n)}x_2^{(n)}\dots x_m^{(n)}$ would be an m-digit decimal greater than $x^{(m)}$ and less than or equal to every element of B. Furthermore, we cannot have $x_i^{(m)} > x_i^{(n)}$. for any i with $0 \leq i \leq m$, since then, with this i, the number $x'' = x_0^{(m)}.x_1^{(m)}x_2^{(n)}\dots x_i^{(m)}00\dots0$ ($n-i$ zeros at the end) would be an n-digit decimal greater than $x^{(n)}$ and less than or equal to every element of B.

Now, the infinite decimal $r = x_0^{(0)}.x_1^{(1)}x_2^{(2)}\dots x_n^{(n)}\dots$ is the number determined by cut (A, B). Indeed, $r \geq x$ for every $x \in A$. Assume, on the contrary, that $r < x$ and let n be such that $10^{-n} < x - r$. Then noting that $r - x^{(n)} < 10^{-n}$, the decimal number $x^{(n)} + 10^{-n} \leq r + 10^{-n} < x$ would be less than every element of B (since x is less than every element of B), contradiction the choice that $x^{(n)}$ was the largest such n-digit decimal fraction. Furthermore, $r \leq y$ for every $y \in B$. Assume, on the contrary that $r > y$ for some $y \in B$, and let n be such that $10^{-n} < r - y$. Then noting that $r - x^{(n)} < 10^{-n}$, we have $x^{(n)} > r - 10^{-n} > y$, contradiction the assumption that $x^{(n)} \leq y$ for every $y \in B$. The proof is complete in case the cut (A, B) is such that A contains a positive number.

If (A, B) is a cut such that B contains the negative number, writing $-A = \{-x : x \in A\}$ $-B = \{-y : y \in B\}$, the cut $(-B, -A)$ is such that $-B$ contains a positive number; hence, by the above argument, this cut determines a number r. Then the cut (A, B) determines the number $-r$.

Finally, if the cut (A, B) is such that A does not contain a positive number and B does not contain a negative number, then it is clear that the cut (A, B) determines the number 0. The proof is complete. □

The Axiom of Completeness guarantees, for example, that the number $\sqrt{2}$ exists. Namely, the cut (A, B) with $A = \{x : x < 0 \text{ or } x^2 \leq 2\}$ and $B = \{x : x > 0 \text{ and } x^2 > 2\}$ and determines the number t such that $t^2 = 2$.

To show this, we will first show that for every $\epsilon > 0$ we have $|t^2 - 2| < \epsilon$. In order to show this, we may assume that $\epsilon < 1$. First note note that there is an $x \in A$ and a $y \in B$ with $x \geq 0$ and $y \geq 0$ (the latter holds for every $y \in B$) such that $y - x \leq \epsilon/6$. To see this, consider the set

$$S = \{n\epsilon/6 : n \geq 0 \text{ is an integer} \quad \text{and} \quad n\epsilon/6 \in A\}.$$

S is not empty, since $0 \in S$. Furthermore, it is clearly a finite set, and so it has a largest element. Now, choose x to be the largest element of this set and put $y = x + \epsilon/6$ (clearly, $y \in B$, since if we had $y \notin B$ then we would have $y \in A$, and so $y \in S$, and then x would not be the largest element of x.

Observe that $x < 2$ (because $x \in A$, and so $x^2 \leq 2$ unless $x < 0$ by the definition of A) and $y = x + \epsilon/6 \leq 2 + 1/6 < 3$ (this is where we used the assumption $\epsilon < 1$). Furthermore, $x \leq t \leq y$, since t is the number determined by the cut (A, B). We have

$$t^2 - 2 \leq t^2 - x^2 = (t - x)(t + x) \leq (y - x)(y + x) < \frac{\epsilon}{6} \cdot 5 < \epsilon;$$

the third inequality holds because we have $y - x \leq \epsilon/6$, $x < 2$, and $y < 3$. Similarly,

$$2 - t^2 \leq y^2 - t^2 = (y - t)(y + t) \leq (y - x)(y + y) < \frac{\epsilon}{6} \cdot 6 = \epsilon.$$

[21.2] i.e., if $x^{(n)} = 345.84$ then $n = 2$, $x_0^{(n)} = 345$, $x_1^{(n)} = 8$, and $x_2^{(n)} = 4$.

[21.3] The set of these decimal fractions is nonempty, since it contains the number 0. It is finite, because if b is an arbitrary element of B, there are at most $10^n \cdot b + 1$ n-digit nonnegative decimal fractions less than or equal to b.

These two inequalities together show that

$$|t^2 - 2| < \epsilon.$$

as claimed.

Since $\epsilon > 0$ was arbitrary, this inequality holds for every $\epsilon > 0$. Now, assume that $t^2 \neq 2$. Then the number $|t^2 - 2|$ is positive. Choosing $\epsilon = |t^2 - 2|$, the above inequality cannot hold. This is a contradiction, showing that we must have $t^2 = 2$.

Let S be a nonempty set of reals. A number t such that $t \geq x$ for every $x \in S$ is called an *upper bound* of S. The set A is called *bounded* from above if it has an upper bound. The number c that is an upper bound of S such that $c \leq t$ for every upper bound t of S is called the *least upper bound* or *supremum* of S. The supremum of S is denoted as $\sup S$.

Lemma 21.1. *Let S be a nonempty set S of reals that is bounded from above. Then S has a supremum.*

Proof. Let B be the set of upper bounds of S, and let A be the set of those reals that are not upper bounds of S (i.e., $A = \mathbb{R} \setminus B$). Then it is clear that (A, B) is a cut. Indeed, if we have $x \geq y$ and $y \in B$ for the reals x and y, then x is also an upper bound of S (since it is at greater than or equal to another upper bound, namely y). This shows that for every $x \in A$ and $y \in B$ we must have $x < y$. A is not empty since S is not empty, so not every real is an upper bound of S; B is not empty, since S is bounded from above by assumption.

Let t be the real determined by the cut (A, B). Then t is an upper bound of S. Assume, on the contrary, that there is an $s \in S$ is such that $s > t$. Let $u = (s + t)/2$. Then $u > t$, so we must have $u \in B$ (since the cut (A, B) determines t). Yet $u < s$, so u is not an upper bound of S, contradicting the relation $u \in B$.

Furthermore, t is the least upper bound of S. Assume, on the contrary, that y is an upper bound of S such that $y < t$. Then $y \in B$ (since B contains every upper bound of S). But then we must have $y \geq t$ (since the cut (A, B) determines t). This contradicts the relation $y < t$. The proof is complete. □

For the empty set, one usually writes that $\sup \emptyset = -\infty$, and for a set S that is not bounded from above, one writes that $\sup S = +\infty$. With this extension, the symbol $\sup S$ will be meaningful for any subset of S reals.

Let S be a nonempty set of reals. A number t such that $t \leq x$ for every $x \in S$ is called a *lower bound* of S. The set A is called *bounded* from below if it has an lower bound. The number c that is an lower bound of S such that $c \geq t$ for every lower bound t of S is called the *greatest lower bound* or *infimum* of S. The infimum of S is denoted as $\inf S$. Every nonempty set that is bounded from below has an infimum. The proof of this statement is similar to the proof of the Lemma above. Instead of carrying out this proof, one can argue more simply that $\inf S = -\sup(-S)$, where $-S \stackrel{def}{=} \{-s : s \in S\}$. One usually writes $\inf \emptyset = +\infty$, and if S is not bounded from below, then one writes $\inf S = -\infty$.

22 Sequences and limits

22.1 Sequences and subsequences

A sequence (of real numbers) is usually thought of as an infinite list numbers; for example

$$1, 1/2, 1/3, 1/4, \ldots, 1/n, \ldots.$$

A more formal way to think of a sequence as a function on the set $\mathbb{N} = \{1, 2, 3, 4 \ldots, m \ldots\}$.[22.1] With this in mind, the above sequence can be thought of as the function $f : \mathbb{N} \to \mathbb{R}$ such that $f(n) = 1/n$. A subsequence of the above sequence is usually thought of as taking only certain members of the sequence; one needs to take an infinite number of members, and one must keep the original order. Thus, a subsequence of the above sequence is

$$1/3, 1/7, 1/12, 1/18, \ldots.$$

With the function description of a sequence, this can formally be described as the composition $f \circ \pi$, where π is an *increasing* (to be defined) function $\pi : \mathbf{N} \to \mathbf{N}$.[22.2] In the given example, we need to have $\pi(1) = 3$, $\pi(2) = 7$, $\pi(3) = 12$, $\pi(4) = 18$, ... A function (on a set of reals or on a set of integers) is called increasing if for any x and y in the domain of f, if $x < y$ then we have $f(x) < f(y)$. The assumption that π in the subsequence $f \circ \pi$ of the sequence f is increasing ensures that the members of f are not rearranged in the subsequence $f \circ \pi$.

One often uses the notation $\{a_n\}$ or $\{a_n\}_{n=1}^{\infty}$ for the sequence described by the function $f(n) = a_n$. This is like writing the argument of the function as a subscript, i.e., writing a_n instead of the function notation $a(n)$. When using the notation $\{a_n\}$ for the sequence, but still does not call the sequence itself a, whereas when one uses the description $f(n) = a_n$, the sequence itself would naturally be the function f, and not $f(n)$ or $\{f(n)\}$.

A subsequence the sequence $\{a_n\}$ is often described as $\{a_{n_k}\}$, where $\{n_k\}$ itself is a sequence. The function description is clearer, and in some contexts is much more useful, even though it is not used often. For example, if we write $f(n) = a_n$ and $\pi(k) = n_k$, we have $f \circ \pi(k) = a_{n_k}$.

22.2 Limits

In an informal discussion, one says that L is the limit of the sequence $\{a_n\}$, in symbols $L = \lim_{n\to\infty} a_n$, if for large n the number a_n is close to L. The problem with this description is that the words "large" and "close" do not have clear mathematical meanings. To correct this deficiency, we will call the number n N-large if $n \geq N$. Here N is usually thought of as a large (whatever that means) integer, but in fact no restriction needs to be put on N, so N need not be large, nor an integer (but it must be a real number) for N-large to make sense. Similarly, for $\epsilon > 0$, we will say that the number L is ϵ-close to the number x if $|L - x| < \epsilon$. Here one usually thinks of ϵ being small (whatever that means), so that ϵ-close really means close, but this again not a requirement for ϵ-close to make sense.

This gives a clue as to how to make the definition of limit precise. What we want is to ensure that a_n is as close to L as we want by making sure that n is large enough. That is, L is called the

[22.1] $\mathbb{N}$ is called the set of natural numbers, hence the symbol used. Here we did not take 0 to be a natural number; in some context, it is better to also include 0. In particular, in discussions of the fundation of mathematics, 0 is also considered a natural number.

[22.2] The letter π is customarily used to denote the ration of the circumference and the diameter of a circle; $\pi \approx 3.1415926\ldots$. Here, of course, π is used in a completely different sense,

limit of the sequence $\{a_n\}$ if for every $\epsilon > 0$ we can find an N such that if n is N-large then a_n is ϵ-close to L.

This is almost the final form, though it assumes that a sequence can only have one limit (this is true, but it is better left as a statement to be proved than to include it in the definition), and then we need to write the meaning of N-large and ϵ-close directly into the definition. That is,

Definition 22.1. We say that L is a limit of the sequence $\{a_n\}$ if for every $\epsilon > 0$ there is an N such that $|L - a_n| < \epsilon$ if $n > N$.

Expressed in a logic formula, we have

$$\lim(\{a_n\}, L) \stackrel{def}{\equiv} (\forall \epsilon > 0)(\exists N)(\forall n > N)(|L - a_n| < \epsilon), \tag{22.1}$$

where ϵ runs over reals, N and n run over integers.[22.3] The symbol $\lim(\{a_n\}, L)$ expresses the assertion that L is a limit of the sequence $\{a_n\}$. We avoided the notation $\lim_{n\to\infty} a_n = L$, since this notation implicitly asserts that a sequence can only have one limit, which is an assertion to be proved. What makes this concept difficult to understand when it is encountered is the alternation of quantifiers, that is, the switch between universal and existential quantifiers. In fact, one way of classifying the complexity of a mathematical formula is based on counting the alternation of quantifiers. It is easy to show that a sequence can only have one limit.

Lemma 22.1. *Let $\{a_n\}$ be a sequence of reals. Then there is at most one L such that L is a limit of this sequence.*

Proof. Assume that L_1 and L_2 are both limits of $\{a_n\}$, where $L_1 \neq L_2$. Let $\epsilon = |L_2 - L_1|/2$. Clearly, $\epsilon > 0$, so there must be N_1 and N_2 such that $|L_1 - a_n| < \epsilon$ for $n \geq N_1$ and $|L_2 - a_n| < \epsilon$ for $n \geq N_2$. Then pick an n such that both $n \geq N_1$ and $n \geq N_2$ hold. Then

$$2\epsilon = |L_2 - L_1| = |(L_2 - a_n) + (a_n - L_1)| \leq |L_2 - a_n| + |a_n - L_1| < \epsilon + \epsilon = 2\epsilon;$$

here, the first inequality holds since $|a + b| \leq |a| + |b|$ for any two reals a and b.[22.4] Comparing the sides, this says that $2\epsilon < 2\epsilon$, which is a contradiction. □

Having shown the uniqueness of limit, we can use the notation $L = \lim_{n\to\infty} a_n$ to indicate that L is the limit of the sequence $\{a_n\}$. A sequence that has a limit is called *convergent*; one that does not have a limit is called *divergent*. A sequence $\{a_n\}_{n=1}^{\infty}$ is called *nondecreasing* if $a_n \leq a_m$ whenever $1 < n < m$.[22.5] The sequence $\{a_n\}_{n=1}^{\infty}$ is called bounded from above if the associated set $\{a_n : n \geq 1\}$ is bounded from above. An upper bound of this set is also called an upper bound of the sequence, and the supremum of this set is called the supremum of the sequence. The supremum of the sequence $\{a_n\}_{n=1}^{\infty}$ is usually denoted as $\sup a_n$ or $\sup_{n\geq 1} a_n$. The words or phrases bounded, bounded from below, supremum, infimum, and the symbol $\inf_{n\geq 1} a_n$ are used for sequences similarly. We have the following

Lemma 22.2. *Let $\{a_n\}_{n=1}^{\infty}$ be a nondecreasing sequence that is bounded from above. Then $\{a_n\}_{n=1}^{\infty}$ has a limit. More precisely, its supremum is also its limit.*

[22.3] While it is customary to require that N be an integer, this is not important. We can allow N run over all reals, but n must run over integers, since a_n is meaningless if n is not a (positive) integer.

[22.4] Indeed, equality holds unless a and b have different signs.

[22.5] This sequence is called *increasing* if $a_n < a_m$ whenever $1 < m < m$. One used to say increasing in the wider sense instead of nondecreasing, and strictly increasing instead of increasing.

Proof. By Lemma 21.1, the secquence $\{a_n\}_{n=1}^{\infty}$ has a supremum; let this be L. Then, given an arbitrary $\epsilon > 0$, $L - \epsilon$ is not an upper bound of the set $\{a_n : n \geq 1\}$. Hence, there is an $N \in \mathbb{N}$ such that $a_N > L - \epsilon$. Since the sequence $\{a_n\}_{n=1}^{\infty}$ is nondecreasing, we have $a_n \geq a_N$ for all $n \geq N$. Hence we have $L - \epsilon < a_n \leq L$ for all $n \geq N$; the second inequality holds since L is an upper bound of the set $\{a_n : n \geq 1\}$. Thus $L - \epsilon < a_n < L + \epsilon$, i.e, $|L - a_n| < \epsilon$, for all $n \geq N$. Since $\epsilon > 0$ was arbitrary, this shows that $\lim_{n\to\infty} a_n = L$. □

22.2.1 A subsequence of a convergent subsequence is convergent

If we have a convergent sequence, then any of its subsequences converges to the same limit. We would like to formulate this rigorously and give a rigorous proof. We recall that, in this rigorous interpretation, a sequence is a function $f : \mathbb{N} \to \mathbb{R}$. Having a proper notation for the sequence, namely f, its limit $\lim_{n\to\infty} f(n)$ can be more conveniently written as $\lim f$; after all, n plays no explicit role here.

Theorem 22.1 (Convergence of a subsequence)**.** *Let $f : \mathbb{N} \to \mathbb{R}$ be a sequence, and let $\pi : \mathbf{N} \to \mathbf{N}$ be an increasing sequence. Assume that $\lim f = L$ for some $L \in \mathbb{R}$. Then $\lim f \circ \pi = L$*

Proof. According to Definition 22.1, $\lim f = L$ means that, given an arbitrary $\epsilon > 0$, there is an integer $N > 0$ such that for every $n > N$, we have $|L - f(n)| < \epsilon$. For such an n, the inequality $\pi(n) \geq n > N$ holds, so we certainly have $\big|L - f\big(\pi(n)\big)\big| < \epsilon$. Hence, again according to Definition 22.1, $\lim f \circ \pi = L$ holds. □

22.3 Limit rules

The well-known rules about limits of functions are also true for limits of sequences. We have

Theorem 22.2. *Let $\{a_n\}_{n=1}^{\infty}$ and $\{b_n\}_{n=1}^{\infty}$ be sequences, and let c be a number, and assume that the limits $A = \lim_{n\to\infty} a_n$ and $B = \lim_{n\to\infty} b_n$ exist. Then the following limits exist and satisfy the equations given:*

$$1)\ \lim_{n\to\infty}(a_n + b_n) = A + B, \qquad 2)\ \lim_{n\to\infty}(a_n - b_n) = A - B, \qquad 3)\ \lim_{n\to\infty} ca_n = cA,$$

$$4)\ \lim_{n\to\infty} a_n b_n = AB, \qquad \textit{and} \qquad 5)\ \lim_{n\to\infty} \frac{a_n}{b_n} = \frac{A}{B},$$

the last one under the additional assumption that $B \neq 0$ and $b_n \neq 0$ for all $n \geq 1$.[22.6]

To follow the proof of this result, some may prefer to think in terms of formula (22.1) rather than the verbal definition preceding it. To simplify telling the proof, we need the following

Lemma 22.3. *Any convergent sequence is bounded.*

Proof. Let L be the limit of the sequence $\{a_n\}_{n=1}^{\infty}$, and let $\epsilon = 1$.[22.7] Using Definition 22.1 or formula (22.1), let N be such that $|L - a_n| < 1$ for $n > N$. We then have

$$|a_n| = \big|(a_n - L) + L\big| \leq |a_n - L| + |L| \leq 1 + |L|$$

[22.6] The requirement that $b_n \neq 0$ can be ignored if one is willing to accept the fact that, without this assumption, finitely many members of the sequence $\{a_n/b_n\}_{n=1}^{\infty}$ may be meaningless, since even then, the limit of the sequence may be meaningfully defined.

[22.7] Any other choice of $\epsilon > 0$ would work.

for all $n > N$. Let M be the maximum of the finite set

$$\{|L| + 1\} \cup \{|a_n| : 1 \leq n \leq N\}.$$

Then $|a_n| \leq M$ for all $n \geq 1$. showing that $\{a_n\}_{n=1}^{\infty}$ is indeed bounded. □

Proof of Theorem 22.2. Let $\epsilon > 0$ be arbitrary. According to Definition 22.1 or formula (22.1), in each of the cases we have to show the existence of an appropriate $N > 0$. In what follows, we will use this definition or this formula repeatedly, without explicit reference.[22.8]

Ad Claim 1) Let N_1 be such that $|A - a_n| < \epsilon/2$ for $n > N_1$; such an N_1 must exist, since for such an N_1 must exist for any positive number (in particular, $\epsilon/2$) replacing ϵ in formula (22.1), since we have $A = \lim_{n\to\infty} a_n$. Similarly, let N_2 be such that we have $|B - b_n| < \epsilon/2$ for $n > N_2$; Writing $N = \max\{N_1, N_2\}$, for $n > N$ we have

$$\big|(a_n + b_n) - (A + B)\big| = \big|(a_n - A) + (b_n - B)\big| \leq |a_n - A| + |b_n - B| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon,$$

establishing the claim.

Before discussing Claim 2), we will fisrt verify Claim 4) and then Claim 3).

Ad Claim 4) According to Lemma 22.3, the sequences $\{a_n\}_{n=1}^{\infty}$ is bounded; let M be such that $|a_n| \leq M$ for all $n \geq 1$. Let N_1 and N_2 be such that[22.9]

$$|A - a_n| < \frac{\epsilon}{2|B| + 1} \text{ for } n > N_1, \quad \text{and} \quad |B - b_n| < \frac{\epsilon}{2M + 1} \text{ for } n > N_2;$$

Let $N = \max\{N_1, N_2\}$. For $n > N$ we have

$$\begin{aligned} |AB - a_n b_n| &= \big|(A - a_n)B + a_n(B - b_n)\big| \leq \big|(A - a_n)B\big| + \big|a_n(B - b_n)\big| \\ &\leq |A - a_n||B| + |a_n||B - b_n| \leq \frac{\epsilon}{2|B| + 1}|B| + M\frac{\epsilon}{2M + 1} \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon. \end{aligned}$$

This establishes the claim.

Ad Claim 3) We will use the already established Claim 4) with $b_n = c$ for all n. Since $\lim_{n\to\infty} c = c$, the assertion follows.

Ad Claim 2) We will use the already establised Claims 1) and 3). We have

$$\begin{aligned} \lim_{n\to\infty}(a_n - b_n) &= \lim_{n\to\infty}(a_n + (-1)b_n) = \lim_{n\to\infty} a_n + \lim_{n\to\infty}(-1)b_n \\ &= \lim_{n\to\infty} a_n + (-1)\lim_{n\to\infty} b_n = A + (-1)B = A - B, \end{aligned}$$

establishing the claim.

Ad Claim 5) We first establish the claim under the assumption that $a_n = 1$; i.e. we will show that under the assumption $B = \lim_{n\to\infty} b_n \neq 0$ we have

$$\lim_{n\to\infty} \frac{1}{b_n} = \frac{1}{B}. \tag{22.2}$$

[22.8]Saying "Ad Claim 1)" next means "Concerning Claim 1)"; "ad" is Latin, meaning "toward," "next to," or something similar. It is commonly used in mathematics when discussing cases.

[22.9]The reason for using $2|B| + 1$ instead of $2|B|$, and $2M + 1$ instead of $2M$ next is to avoid having a zero in the denominator, since M, A, or B might be zero.

To this end, let N be such that

$$|B - b_n| < \min\left\{\frac{|B|}{2}, \frac{B^2\epsilon}{2}\right\} \tag{22.3}$$

for $n > N$. First observe that the inequality $|B - b_n| < |B|/2$ holds for all $n > N$; this implies that $|b_n| > |B|/2$ for $n > N$. Indeed, assuming, on the contrary, that $|b_n| \leq |B|/2$ for some $n > N$, for such an n we have

$$|B| = \big|(B - b_n) + b_n\big| \leq |B - b_n| + |b_n| < \frac{|B|}{2} + \frac{|B|}{2} = |B|,$$

which is a contradiction, since it says $|B| < |B|$; note that strict inequality holds since the inequality $|B - b_n| < |B|/2$ is strict. Hence, for $n > N$ we have

$$\left|\frac{1}{B} - \frac{1}{b_n}\right| = \left|\frac{b_n - B}{Bb_n}\right| < \left|\frac{b_n - B}{B \cdot (B/2)}\right| = \left|\frac{2(b_n - B)}{B^2}\right| < \epsilon,$$

where first inequality holds since $|b_n| > |B|/2$, and the second one holds in view of the second element of the mininum listed on right-hand side of inequality (22.3). Therefore, equation (22.2) follows.

With this Claim 5) easily follows by using the already established Claim 4). Indeed, we have

$$\lim_{n\to\infty} \frac{a_n}{b_n} = \lim_{n\to\infty} a_n \frac{1}{b_n} = \lim_{n\to\infty} a_n \lim_{n\to\infty} \frac{1}{b_n} = A\,\frac{1}{B} = \frac{A}{B},$$

establishing Claim 5). □

23 Supremum and limits

23.1 Closed and open sets

Lemma 23.1. *Let S be a nonempty set of reals that is bounded from above, and let $a = \sup S$. Let $\epsilon > 0$. Then the interval $(a - \epsilon, a]$ contains an element of S.*

Proof. As $a - \epsilon$ is not an upper bound of S (a being its least upper bound), there must be an $x \in S$ with $x > a - \epsilon$. We have $x \leq a$, since a is an upper bound of S. Thus $x \in S \cap (a - \epsilon, a]$. □

Definition 23.1. A set S of reals is said to be open if for every $x \in S$ there is an $\epsilon > 0$ such that the interval $(x - \epsilon, x + \epsilon)$ is a subset of S.

What is meant in the above definition is "if and only if" (that is, "S is said to be open *if and only if* for every ..."). In mathematical definitions, it is customary to say *if* in similar cases when one means *if and only if*. In other situations in mathematics, one makes a very careful distinction between "if" and "if and only if."

Given a set $S \subset \mathbb{R}$, we will write

$$\complement S \stackrel{def}{=} \mathbb{R} \setminus S, \tag{23.1}$$

and call it the complement of S.

Definition 23.2. A set S of reals is said to be closed if its complement $\complement S$ is open.

Lemma 23.2. *Let S be a nonempty closed set of reals that is bounded from above. Then S has a maximum. In other words, there is a $u \in S$ such that $u \geq s$ for every $s \in S$.*

Proof. As S is nonempty and bounded from above, it has a supremum; write $u = \sup S$. We will prove that $u \in S$; then it will be clear that $u = \max S$ (i.e. that $u \in S$ and $x \leq u$ for every $x \in S$).

Assume, on the contrary, that $u \notin S$. Then $u \in \complement S$. As S is closed, its complement $\complement S$ is open; thus, there is an interval $(u - \epsilon, u + \epsilon)$ with center u that is included in $\complement S$ ($\epsilon > 0$); Hence S has no elements in this interval; since u is an upper bound of S (i.e., $x \leq u$ for every $x \in S$), this means that $u - \epsilon$ is an upper bound of S (i.e., $x \leq u - \epsilon$ for every $x \in S$). This contradicts the assumption that u is the least upper bound of S. Therefore the assumption $u \notin S$ must be wrong, completing the proof. □

Lemma 23.3. *Let S be a set of reals that is not closed. Then there is a real number $x \notin S$ and a sequence $\{a_n\}_{n=1}^{\infty}$ of elements of S that converges to x.*

Proof. As S is not closed, its complement $\complement S$ is not open. Therefore, there is a real $x \in \complement S$ such that no no open interval with center x is included in $\complement S$, that is, for every $\epsilon > 0$ the set

$$(x - \epsilon, x + \epsilon) \cap S$$

is not empty (if it were, the interval on the left would be included in $\complement S$). Using this statement with $\epsilon = 1/n$, for each positive integer n we can find a a_n in this set; that is $a_n \in S$ and $|x - a_n| < 1/n$. From the latter inequality we can conclude that $\lim_{n\to\infty} a_n = x$. Thus the sequence $\{a_n\}_{n=1}^{\infty}$ has the desired properties. □

23.2 Uses of the Axiom of Replacement and of the Axiom of Choice

The proof of Lemma 23.3 is an interesting interplay of the Axiom of Replacement (see formula (20.2) and the Axiom of Choice (see formula (20.1)). Namely, the way of picking a_n as described in the proof cannot be done in a formal proof, since it is an infinite process, and a formal proof as needs to be a finite sequence of logic formulas as described in Subsection 6.1. The way out from this difficulty is to take a choice function f on the set $\mathcal{P}(S) \setminus \emptyset$. Then we can define a formula $\Phi(u, v)$ such that for all integers $n \in \mathbb{N} = \{1, 2, 3, \ldots\}$ we have $\Phi(n, v)$ if and only if $y = \Big(n, f\big((x - 1/n, x + 1/n) \cap S\big)\Big)$. Here x and S are as given in Lemma 23.3. Then one can define the sequence $g = \{a_n\}_{n=1}^{\infty}$ (i.e., $g : \mathbb{N} \to \mathbb{R}$ is such that $g(n) = a_n$ for all n) according to the Axiom of Replacement by the formula

$$g = \{v : u \in \mathbb{N} \quad \textit{and} \quad \Phi(u, v)\}.$$

This proof can be carried out with a much weeker version of the Axiom of Choice than described above, since the choice function f needs to be defined only on the countable set

$$\{(x - 1/n, x + 1/n) \cap S : n \in \mathbb{N}\}.$$

The weakened version of the Axiom of Choice that asserts the existence of a choice function only for countable sets is usually called the Axiom of Countable Choice. Countable Choice is not sufficient for many arguments in analysis. A stronger axiom is the Axiom of Dependent Choice, which still only makes countably many choices. This says that given a nonempty set A and a relation R on A such that for all $x \in A$ there is a $y \in A$ such that xRy, there for every $u \in R$ there is a function $f : \mathbb{N} \to A$ such that $f(q) = u$ and for all $n \in \mathbb{N}$ we have $f(n)Rf(n + 1)$.

23.3 More on supremum and limits

Lemma 23.4. *Let a be such that $0 < a < 1$. Then $\lim_{n\to\infty} a^n = 0$.*

Proof. The assumptions imply that $0 < a^{n+1} < a^n$; thus the sequence $\{a^n\}_{n=1}^{\infty}$ is decreasing and bounded from below; therefore it has a limit according to Lemma 22.2. Write $x = \lim_{n\to\infty} a^n$. Clearly, we also have $x = \lim_{n\to\infty} a^{n+1}$. This can easily be verified directly. Alternatively, one may observe that $\{a^{n+1}\}_{n=1}^{\infty}$ is a subsequence[23.1] of $\{a^n\}_{n=1}^{\infty}$; hence, the former sequence must have the same limit as the latter. Now,

$$ax = a \cdot \lim_{n\to\infty} a^n = \lim_{n\to\infty} a \cdot a^n = \lim_{n\to\infty} a^{n+1} = x,$$

i.e., $ax = x$, or $(a-1)x = 0$. As $a \neq 1$ by our assumptions, this is only possible if $x = 0$. Thus, $\lim_{n\to\infty} a^n = 0$, which is what we wanted to prove. □

Second proof. The assertion $\lim_{n\to\infty} a^n = 0$ can also be proved by using Bernoulli's Inequality, saying that $(1+x)^n \geq 1 + nx$ holds whenever $x \geq -1$, for every positive integer n (see [16, formula (1) on p. 1]). For $x \geq 0$, Bernoulli's Inequality is a direct consequence of the Binomial Theorem. The case $-1 \leq x < 0$ of Bernoulli's Inequality is harder to establish, but this case is not needed for proving $\lim_{n\to\infty} a^n = 0$.

Writing $x = 1/a - 1$, we have $x \geq 0$ and $a = 1/(1+x)$, and so $(1/a)^n = (1+x)^n \geq 1 + nx$, i.e., $a^n \leq 1/(1+nx)$. Given $\epsilon > 0$, we will have $1/(1+nx) < 1/(nx) \leq \epsilon$ whenever $n \geq N = 1/(x\epsilon)$. For such n we will have $-\epsilon < a^n - 0 < \epsilon$, showing that $\lim_{n\to\infty} a^n = 0$. □

Lemma 23.5. *Let A and B be two nonempty sets of reals that are bounded from above. Then*

$$\sup\{x + y : x \in A \text{ and } y \in B\} = \sup A + \sup B.$$

Proof. Writing $S = \sup\{x + y : x \in A \text{ and } y \in B\}$, $a = \sup A$, and $b = \sup B$, we will first show that $\sup S \leq a + b$. To this end, let $x + y$ be an element of S, where $x \in A$ and $y \in B$. Then $x \leq a$ (since a is an upper bound of A; in fact, it is its least upper bound) and $y \leq b$ (since b is an upper bound of B). Thus $x + y \leq a + b$. As $x + y$ was an arbitrary element of S, this shows that $a + b$ is an upper bound of S. Therefore $\sup S \leq a + b$, since $\sup S$ is the *least* upper bound of S.

Next we will show that $\sup S \geq a + b$. We will do this by showing that no number $c < a + b$ is an upper bound of S. So, let $c < a + b$ be arbitrary and write $\epsilon = (a + b - c)/2$; then $\epsilon > 0$. Now $a - \epsilon$ is not an upper bound of A (as a is its least upper bound), so there must be an $x \in A$ with $x > a - \epsilon$. Similarly, $b - \epsilon$ is not an upper bound of B, so there must be a $y \in B$ with $y > b - \epsilon$. Then

$$x + y > (a - \epsilon) + (b - \epsilon) = a + b - 2\epsilon = c.$$

As $x + y \in S$, this implies that c is not an upper bound of S, as we wanted to show. As both $\sup S \leq a + b$ and $\sup S \geq a + b$ hold, we must have $\sup S = a + b$. The proof is complete. □

Second proof. Using the notation introduced in the first proof, we will give a second proof of the inequality $\sup S \geq a + b$ (we will not give another proof of the inequality $\sup S \leq a + b$; the proof of this latter inequality will have to be taken from the first solution). Let c be an upper bound of S; it will be enough to show that $c \geq a + b$.

[23.1] The sequence $\{b_n\}_{n=1}^{\infty}$ is called a subsequence of $\{a_n\}_{n=1}^{\infty}$ if $b_n = a_{f(n)}$ for some strictly increasing function f that is defined for each positive integer and has positive integers as values. In the present case, one can put $a_n = a^n$ and $b_n = a^{n+1} = a_{n+1}$, that is, $f(n) = n + 1$.

Let $x \in A$ and $y \in B$ be arbitrary. Then $x + y \in S$, and so $x + y \leq c$, since c is an upper bound of S. That is, $x \leq c - y$. Now, consider this inequality for a fixed $y \in B$. Then we can see that, for every $x \in A$, the inequality $x \leq c - y$ holds. I.e., $c - y$ is an upper bound of A, and so $c - y \geq a$, the least upper bound of A.

The last inequality can also be written as $y \leq c - a$; this inequality holds for every element $y \in B$ (since $y \in B$ was arbitrary). Thus $c - a$ is an upper bound of B; i.e., $c - a \geq b$, since b is the least upper bound of this set. Thus $c \geq a + b$, as we wanted to show. □

24 Limits of functions

24.1 The precise definition

In an informal definition of limit of function, one would say that L is the limit of the function f at a if, whenever x is close to a but $x \neq a$, the function value $f(x)$ is close to L. The problem with this definition is similar to the problem with the informal definition of sequences given in Subsection 22.2: the term "close" is does not have a clear mathematical meaning. We already described there how to correct this deficiency. For $\epsilon > 0$, we will say that the number L is ϵ-close to the number x if $|L - x| < \epsilon$. Here one usually thinks of ϵ being small (whatever that means), so that ϵ-close really means close, but this again not a requirement for ϵ-close to make sense.

To use this concept to further clarify the meaning of the limit of a function. For L to be the limit of f at a, we want to ensure that $f(x)$ is arbitrarily close to L provided that we make x close enough to a ($x \neq a$). That is, for every $\epsilon > 0$, we can ensure that $f(x)$ is ϵ-close to L as long as we can find a $\delta > 0$ such that $x \neq a$ is δ-close to a. This now can be described in a logic formula:

$$(\forall \epsilon > 0)(\exists \delta > 0)(\forall x)\Big(\big(|x - a| < \delta \,\&\, x \neq a\big) \to \big|L - f(x)\big| < \epsilon\Big).$$

There are a couple of things to tweak about this formula. First, the subformula $|x - a| < \delta \,\&\, x \neq a$ can be more conveniently expressed as $0 < |x - a| < \delta$. Secondly, if x is not in the domain of $f(x)$, the question if whether to consider $|L - f(x)| < \epsilon$ true if $x \notin \mathrm{dom}(f)$; we will consider it false.

To clarify this issue, we need to translate the statement $|L - f(x)| < \epsilon$ into *primitive*[24.1] terms. If $f(x)$ is defined, then this statement can be translated in two different ways in terms of the discussion of functions given in Section 18: $(\forall y)\big((x, y) \in f \to |L - y| < \epsilon\big)$ and $(\exists y)\big((x, y) \in f \,\&\, |L - y| < \epsilon\big)$. The two translations give the same result if $x \in \mathrm{dom}(f)$. If $x \notin \mathrm{dom}(f)$ then the first translation is always true (a conditional with false antecedent is true), and the second one is always false. For this reason, we need to take the second translation.

One often considers one-sided limits, and it would be nice to set up the definition in a way that this is asserted in the definition. This can be done by defining the limit for the case when x approached a through elements of a given set S; to describe limit from the left, we would take $S = (-\infty, a]$, and to describe limit from the right, we would take $S = [a, +\infty)$ Using the formula $\lim(f, a, S, L)$ to say that L is a limit of f when x approaches a through elements of the set S, we

[24.1] I.e., original; that is in terms of the definition of functions described in Section 18. In mathematics, "primitive" is not a derogatory term. For example, axiomatic set theory uses two primitive symbols: $=$ and $\in$. This means that these symbols will not be defined, it will not be explained in the formal system what they are about, but they need to satisfy certain axioms. When one says that a set is a collection, that belongs to the intuitive description and not to the formal system. Another example comes from calculus: an antiderivative is also called primitive function.

can write

$$\begin{aligned}&\lim(f,a,S,L) \stackrel{def}{\equiv}\\ &\qquad(\forall\epsilon>0)(\exists\delta>0)\big(\forall x\in S\big)\Big(0<|x-a|<\delta\to\big|L-f(x)\big|<\epsilon\Big).\end{aligned}\tag{24.1}$$

In this formula, the inequality $\big|L-f(x)\big|<\epsilon$ should be interpreted as false if $x\notin\mathrm{dom}(f)$. Using the traslation of the last inequality in this formula given above, this be written as

$$\begin{aligned}&\lim(f,a,S,L) \stackrel{def}{\equiv}\\ &\qquad(\forall\epsilon>0)(\exists\delta>0)\big(\forall x\in S\big)\Big(0<|x-a|<\delta\to(\exists y)\Big((x,y)\in f\,\&\big|L-y)\big|<\epsilon\Big)\Big).\end{aligned}\tag{24.2}$$

We can recast this definition in words:

Definition 24.1 (Limit of a function)**.** Let f be a real valued function defined on a set of reals, $a,L\in\mathbb{R}$, let $S\subset\mathbb{R}$. We say that L is a limit of f at a in the set S if for every $\epsilon>0$ there is a $\delta>0$ such that for all $x\in S$, if $0<|x-a|<\delta$ then $f(x)$ is defined and $\big|L-f(x)\big|<\epsilon$.[24.2]

24.2 When a limit is not unique

The question of uniqueness of limits is somewhat more complicated than in case of sequences. This is because, given $\delta>0$, there might not exist an $x\in S$ for which $0<|x-a|<\delta$. If this is the case, the formula is vacuously true for any L; that is, every number L is a limit of f at a in S.

To see why this is true, we can rewrite the restricted quantifies $\forall x\in S$ as follows (see Subsection (3.7):

$$\begin{aligned}&\big(\forall x\in S\big)\Big(0<|x-a|<\delta\to\big|L-f(x)\big|<\epsilon\Big)\\ &\qquad\equiv(\forall x)\Big(x\in S\to\Big(0<|x-a|<\delta\to\big|L-f(x)\big|<\epsilon\Big)\Big)\\ &\qquad\equiv(\forall x)\Big(\big(x\in S\,\&\,0<|x-a|<\delta\big)\to\big|L-f(x)\big|<\epsilon\Big);\end{aligned}\tag{24.3}$$

the second equivalence follows from the important tautology

$$\big(A\to(B\to C)\big)\leftrightarrow\big((A\,\&\,B)\to C\big).$$

Now, if there is no x for which

$$x\in S\,\&\,0<|x-a|<\delta$$

holds, then the formula on the right-hand side of equation (24.3) is vacuously true, since the antecedent in the conditional in te last displayed formula is false for a small enough δ.[24.3]

24.2.1 Cluster points of a set

Points where this does not happen are called cluster points:

Definition 24.2 (Cluster point)**.** Let a be a real number and let S be a set of reals. The number a is called a cluster point of S if for every $\delta>0$ there is an $x\in S$ such that $0<|x-a|<\delta$.

In this definition, it is unimportant whether or not we have $a\in S$.

[24.2] Note that $a\in S$ is not required.

[24.3] That is, there is a $\delta>0$ for which this antecedent is false.

24.3 Uniqueness of limits

The point a being a cluster point of the set S guarantees the uniqueness of limit:

Lemma 24.1. *Let a be a real number, f a real-valued function on a set of reals, $S \subset \mathbb{R}$ a set, and $a \in \mathbb{R}$ a number. Assume that a is a cluster point of the set S. Then there is at most one number L that is the limit of f at a in S, i.e., such that* $\lim(f, a, S, L)$.

Proof. Assume that L_1 and L_2 are both limits of f at a in S, where $L_1 \neq L_2$. Let $\epsilon = |L_2 - L_1|/2$. Clearly, $\epsilon > 0$, so there must be δ_1 and δ_2 such that $|L_1 - f(x)| < \epsilon$ for all $x \in S$ with $0 < |x - a| < \delta_1$, and $|L_2 - f(x)| < \epsilon$ for all $x \in S$ with $0 < |x - a| < \delta_2$. Let $\delta = \min\{\delta_1, \delta_2\}$. As a is a cluster point of the set S, there is an $x \in S$ such that $0 < |x - a| < \delta$. For such an x we have

$$2\epsilon = |L_2 - L_1| = \Big|\big(L_2 - f(x)\big) + \big(f(x) - L_1\big)\Big| \leq \big|L_2 - f(x)\big| + \big|f(x) - L_1\big| < \epsilon + \epsilon = 2\epsilon.$$

Comparing the sides, this says that $2\epsilon < 2\epsilon$, which is a contradiction. □

On account of this lemma, if a is a cluster point of the set S, we can use the symbol $\lim_{x \in S: x \to a} f(x) = L$ instead of $\lim(f, a, S, L)$. If we also have $S = \mathbb{R}$, we can use the even simpler notation $\lim_{x \to a} f(x) = L$.

24.4 Limit rules

Given functions f and g on subsets of $\mathbb{R}$ into $\mathbb{R}$, the functions $f \pm g$, fg can be defined on the set $\mathrm{dom}(f) \cap \mathrm{dom}\, g$, cf for a constant c can defined on the set $\mathrm{dom}(f)$, and the function f/g can be defined on the set $\{x \in \mathrm{dom}(f) \cap \mathrm{dom}(g) : g(x) \neq 0\}$.[24.4] We have

Theorem 24.1. *Let a and c be a real numbers, f and g be real-valued functions on sets of reals, let $S \subset \mathbb{R}$ be a set. Assume that a is a cluster point of the set S. Assume further that the limits $A = \lim_{x \in S: x \to a} f(x)$ and $B = \lim_{x \in S: x \to a} g(x)$ exist. Then the following limits exist and satisfy the equations given:*

$$1) \lim_{x \in S: x \to a} \big(f(x) + g(x)\big) = A + B, \qquad 2) \lim_{x \in S: x \to a} \big(f(x) - g(x)\big) = A - B,$$

$$3) \lim_{x \in S: x \to a} cf(x) = cA, \qquad 4) \lim_{x \in S: x \to a} f(x)g(x) = AB,$$

$$\textit{and} \qquad 5) \lim_{x \in S: x \to a} \frac{f(x)}{g(x)} = \frac{A}{B},$$

In the proof, all mathematical relations involving $f(x)$ in case $f(x)$ is not defined should be evaluated as false.[24.5] To follow the proof of this result, some may prefer to think in terms of formula (24.1 the verbal definition following it. Before we start out with the proof, we need a replacement for Lemma 22.3.

[24.4] There is some danger in using these operations without assuming that $\mathrm{dom}(f) = \mathrm{dom}(g)$. For example, the domain of $f + g$ is $\mathrm{dom}(f) \cap \mathrm{dom}(g)$. This is also the domain of $(f + g) - g$, and this is not necessarily the same as $\mathrm{dom}(f)$. So the equation $(f + g) - g = f$ might fail in that the domains of the left-hand side and the right-hand side might not be the same. This will of course not happen if we assume $\mathrm{dom}(f) = \mathrm{dom}(g)$.

Yet this occurs in practice all the time. This of the example $F(x) = f(x) + g(x)$, where $f(x) = x$ and $g(x) = 1/x$, each function having domain where the equation defining it is meaningful. Or think of the solution formula of the quadratic equation as a function of the coefficients.

[24.5] For example, $f(x) > 1$ is translated as $(\exists y)\big((x, y) \in f \,\&\, y > 1\big)$. On the other hand $f(x) \leq 1$ is translated as as $(\exists y)\big((x, y) \in f \,\&\, y \leq 1\big)$. If $f(x)$ is not defined, both are false. On the other hand, the formula $\neg\big(f(x) > 1\big)$ is true. This requires some care, since one would expect that $f(x) \leq 1$ and $\neg\big(f(x) > 1\big)$ mean the same thing; this is only true if $f(x)$ is defined.

Lemma 24.2. *Let a be a real number, f a real-valued function on a set of reals, $S \subset \mathbb{R}$ a set, and $a \in \mathbb{R}$ a number. Assume that a is a cluster point of the set S. Assume $L = \lim_{x \in S: x \to a} f(x)$. Then there is a $\delta > 0$ such that*

$$|f(x)| < |L| + 1 \qquad \text{for all } x \in S \text{ with } 0 < |x - a| < \delta. \tag{24.4}$$

The assumption on x being a cluster point of the set $S \subset \mathbb{R}$ is superfluous here: if it is not satisfied, then the assertion of the lemma is vacous, since we can take a small enough δ for which there is no x as described in formula (24.4).

Proof. and let $\epsilon = 1$.Using Definition 24.1 or formula (24.1) let $\delta > 0$ be be such that $|L - f(x)| < 1$ for $x \in S$ with $0 < |x - a| < \delta$. We then have

$$\big|f(x)\big| = \Big|\big(f(x) - L\big) + L\Big| \leq \big|f(x) - L\big| + |L| \leq 1 + |L|$$

for all $x \in S$ with $0 < |x - a| < \delta$, establishing the lemma. □

Proof. Let $\epsilon > 0$ be arbitrary. According to Definition 24.1 or formula (24.1), in each of the cases we have to show the existence of an appropriate $\delta > 0$. In what follows, we will use this definition or this formula repeatedly, without explicit reference.

Ad Claim 1) Let δ_1 be such that $|A - f(x)| < \epsilon/2$ for $x \in S$ with $0 < |x - a| < \delta_1$; such a δ_1 must exist, since for such a δ_1 must exist for any positive number (in particular, $\epsilon/2$) replacing ϵ in formula (24.1), since we have $A = \lim_{x \in S: x \to a} f(x)$. Similarly, let δ_2 be such that we have $|B - g(x)| < \epsilon/2$ for $x \in S$ with $0 < |x - a| < \delta_2$. Writing $\delta = \min\{\delta_1, \delta_2\}$, let x be such that $x \in S$ with $0 < |x - a| < \delta$; such an x exists by the assumption that a is a cluster point of the set S. We have

$$\begin{aligned}\Big|\big(f(x) + g(x)\big) - (A + B)\Big| &= \Big|\big(f(x) - A\big) + \big(g(x) - B\big)\Big| \\ &\leq \big|f(x) - A\big| + \big|g(x) - B\big| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon,\end{aligned}$$

establishing the claim.

Before discussing Claim 2), we will fisrt verify Claim 4) and then Claim 3).

Ad Claim 4) According to Lemma 24.2 there is a $\delta_0 > 0$ such that $|f(x)| < |A| + 1$ for all $x \in S$ with $0 < |x - a| < \delta_0$. Let $\delta_1 > 0$ be such that[24.6]

$$|A - f(x)| < \frac{\epsilon}{2|B| + 1} \qquad \text{for } x \in S \text{ with } 0 < |x - a| < \delta_1.$$

Let δ_2 be such that

$$|B - g(x)| < \frac{\epsilon}{2|A| + 2} \qquad \text{for } x \in S \text{ with } 0 < |x - a| < \delta_2.$$

Let $\delta = \min\{\delta_0, \delta_1, \delta_2\}$. Let $x \in S$ be such that $0 < |x - a| < \delta$; there is such an x, because a is a cluster point of this set. We have

$$\begin{aligned}\big|AB - f(x)g(x)\big| &= \Big|\big(A - f(x)\big)B + f(x)\big(B - g(x)\big)\Big| \leq \big|(A - f(x))B\big| + \Big|f(x)\big(B - g(x)\big)\Big| \\ &\leq \big|A - f(x)\big||B| + \big|f(x)\big|\big|B - g(x)\big| \leq \frac{\epsilon}{2|B| + 1}|B| + (|A| + 1)\frac{\epsilon}{2|A| + 2} \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.\end{aligned}$$

[24.6]The reason for using $2|B| + 1$ instead of $2|B|$ next is to avoid having a zero in the denominator, since B might be zero.

This establishes the claim.

Ad Claim 3) We will use the already established Claim 4) with $g(x) = c$ for all x. Since $\lim_{x\to a} c = c$, the assertion follows.

Ad Claim 2) We will use the already estabilised Claims 1) and 3). We have

$$\begin{aligned}\lim_{x\in S:x\to a}\big(f(x)-g(x)\big) &= \lim_{x\in S:x\to a}\big(f(x)+(-1)g(x)\big) = \lim_{x\in S:x\to a} f(x) + \lim_{x\in S:x\to a}(-1)g(x)\\ &= \lim_{x\in S:x\to a} f(x) + (-1)\lim_{x\in S:x\to a} g(x) = A + (-1)B = A - B,\end{aligned}$$

establishing the claim.

Ad Claim 5) We first establish the claim under the assumption that $f(x) = 1$ for all x; i.e. we will show that under the assumption $B = \lim_{x\in S:x\to a} g(x) \neq 0$ we have

$$\lim_{x\in S:x\to a} \frac{1}{g(x)} = \frac{1}{B}. \tag{24.5}$$

To this end, let $\delta > 0$ be such that

$$\big|B - g(x)\big| < \min\left\{\frac{|B|}{2}, \frac{B^2\epsilon}{2}\right\}. \tag{24.6}$$

for $x \in S$ with $0 < |x - a| < \delta$. First observe that the inequality $|B - g(x_| < |B|/2$ for such an x, since $\epsilon < |B|/2$; this implies that $|g(x)| > |B|/2$ for such an x. Indeed, assuming, on the contrary, that we have $|g(x)| \leq |B|/2$ for some $x \in S$ with $0 < |x - a| < \delta$, for such an x we have

$$|B| = \Big|\big(B - g(x)\big) + g(x)\Big| \leq \big|B - g(x)\big| + \big|g(x)\big| < \frac{|B|}{2} + \frac{|B|}{2} = |B|,$$

which is a contradiction, since it says $|B| < |B|$; note that strict inequality holds since the inequality $|B - g(x)| < |B|/2$ is strict. Hence, for all $x \in S$ with $0 < |x - a| < \delta$ we have

$$\left|\frac{1}{B} - \frac{1}{g(x)}\right| = \left|\frac{g(x) - B}{Bg(x)}\right| < \left|\frac{g(x) - B}{B\cdot(B/2)}\right| = \left|\frac{2(g(x) - B)}{B^2}\right| < \epsilon,$$

where first inequality holds since $\big|g(x)\big| > |B|/2$, and the second one holds in view of the second element of the mininum listed on right-hand side of inequality (24.6). Therefore, equation (24.5) follows.

With this Claim 5) easily follows by using the already established Claim 4). Indeed, we have

$$\lim_{x\in S:x\to a} \frac{f(x)}{g(x)} = \lim_{x\in S:x\to a} f(x)\,\frac{1}{g(x)} = \lim_{x\in S:x\to a} f(x)\,\lim_{x\in S:x\to a}\frac{1}{g(x)} = A\,\frac{1}{B} = \frac{A}{B},$$

establishing Claim 5). □

References

[1] Carl Boyer. *A History of Mathematics*. Wile International Edition, New York, 1968. `https://archive.org/details/AHistoryOfMathematics`.

[2] Gary Chartrand, Albert D. Polimeni, and Ping Zhang. *Mathematical Proofs: A Transition to Advanced Mathematics*. Pearson, New York, NJ, fourth edition, 2018.

[3] Theodor Estermann. The irrationality of $\sqrt{2}$. *Math. Gaz.*, 59(408):110, 1975. Stable URL: http://www.jstor.org/stable/3616647.

[4] K. Gödel. *The Consistency of Axiom of Chocie and of the Generalized Continuum-hyposthesis with the Axioms of Set Theory*. Princeton University Press, Princeton, NJ, 1940.

[5] Colin Richard Hughes. Irrational roots. *Math. Gaz.*, 83(498):502–503, 1999. Stable URL: http://www.jstor.org/stable/3620971.

[6] L. J. Lander and T. R. Parkin. Counterexample to Euler's conjecture on sums of like powers. *Bull. Amer. Math. Soc.*, 51 (184):1079, 1966.

[7] Attila Máté. The 2003 Brooklyn College Mathematics Prize Exam. http://www.sci.brooklyn.cuny.edu/ mate/prize/2003/all.pdf, 2003.

[8] Attila Máté. The 2004 Brooklyn College Mathematics Prize Exam. http://www.sci.brooklyn.cuny.edu/ mate/prize/2004/all.pdf, 2004.

[9] Attila Máté. The 2006 Brooklyn College Mathematics Prize Exam. http://www.sci.brooklyn.cuny.edu/ mate/prize/2006/all.pdf, 2006.

[10] Attila Máté. The 2010 Brooklyn College Mathematics Prize Exam. http://www.sci.brooklyn.cuny.edu/ mate/prize/2010/all.pdf, 2010.

[11] Attila Máté. The 2011 Brooklyn College Mathematics Prize Exam. http://www.sci.brooklyn.cuny.edu/ mate/prize/2011/all.pdf, 2011.

[12] Attila Máté. The 2012 Brooklyn College Mathematics Prize Exam. http://www.sci.brooklyn.cuny.edu/ mate/prize/2012/all.pdf, 2012.

[13] Attila Máté. The 2013 Brooklyn College Mathematics Prize Exam. http://www.sci.brooklyn.cuny.edu/ mate/prize/2013/all.pdf, 2013.

[14] Attila Máté. The 2014 Brooklyn College Mathematics Prize Exam. http://www.sci.brooklyn.cuny.edu/ mate/prize/2014/all.pdf, 2014.

[15] Attila Máté. The 2015 Brooklyn College Mathematics Prize Exam. http://www.sci.brooklyn.cuny.edu/ mate/prize/2015/all.pdf, 2015.

[16] Attila Máté. The natural exponential function. http://www.sci.brooklyn.cuny.edu/~mate/misc/exp_x.pdf, September 2015.

[17] Attila Máté. The 2016 Brooklyn College Mathematics Prize Exam. http://www.sci.brooklyn.cuny.edu/ mate/prize/2016/all.pdf, 2016.

[18] Attila Máté. The 2017 Brooklyn College Mathematics Prize Exam. http://www.sci.brooklyn.cuny.edu/ mate/prize/2017/all.pdf, 2017.

[19] Attila Máté. The 2018 Brooklyn College Mathematics Prize Exam. http://www.sci.brooklyn.cuny.edu/ mate/prize/2018/all.pdf, 2018.

[20] Attila Máté. Telescoping sums. http://www.sci.brooklyn.cuny.edu/~mate/misc/telescoping_sums.pdf, October 2019.

[21] Attila Máté. The 2020 Brooklyn College Mathematics Prize Exam. `http://www.sci.brooklyn.cuny.edu/ mate/prize/2020/all.pdf`, 2020.

[22] George Polya. *How to Solve It: A New Aspect of Mathematical Method*. Princeton Science Library. Princeton University Press, Princeton, NJ, Princeton Science Li edition, 2015. First edition published in 1945.

[23] W. V. Quine. *Mathematical Logic*. Harvard University Press, Cambridge, MA, USA, revised edition, 1991.

[24] Daniel J. Velleman. *How to prove it: A Structured Approach*. Canbrudge University Press, Cambridge, UK, 3rd edition, 2019.