Studying Mathematics at University

Oscar de Wit

Bachelor Student, Radboud University, Nijmegen, the Netherlands

December 27, 2018

Contents

Preface

I have written this article as a response to my own question when I was a high school student, which was: 'Where can I find a good source to learn more about mathematics as it is taught at universities?'. And so, this article is aimed at high school students with the same question and people in general who want to know what studying mathematics at a university means. So this can be considered an introduction to mathematics at universities.

Supposed knowledge

The following subjects are almost in every high school curriculum. If you do not know about the following subjects, you can learn about them on your own. They are not difficult and easy to learn by the internet. I will refer to some concepts from these topics in this book.

1. Simple algebra (solving linear equations, solving real quadratic equations, factorial)
2. Simple number theory (modular arithmetic, even/uneven arguments)
3. Beginner Calculus (differentiation, integration, sum notation, optimization, trigonometry, absolute value)
4. Complex numbers (Euler's formula, conjugation, complex roots)
5. Triangle inequality for real numbers
6. Simple 3D/2D geometry (reflections, projections)
7. Some sources are:
   (a) 3Blue1Brown at YouTube for videos `http://www.3blue1brown.com/`
   (b) Paul's Online Math Notes at `http://tutorial.math.lamar.edu/`
   (c) `https://en.wikipedia.org/wiki/Portal:Mathematics`
   (d) `https://www.khanacademy.org/math`
8. *How to Prove It: A Structured Approach*, D. J. Velleman, would be a good alternative to the 'Proving theorems' section in this book.

Introduction

This book is divided into a first part where I will introduce some basic concepts which will definitely recur in every mathematics course at a university. This is to provide one with the basic tools. After that there will be some definitions of some common concepts in areas of mathematics and some theorems in these areas. Along the way there will also be exercises to foster thought about the content of the texts. Next to that, I would like to give an idea of what this mathematics is about. For mathematics can be divided into many different subjects which sometimes overlap in a very elegant way.

Furthermore, it is very important, in the first stages of learning advanced mathematics, to see for yourself the distinction between *real formal*, *formal*, *right informal* and *wrong informal* mathematics. Real formal mathematics is in the language of general mathematical logic, which is a complete subject on itself, which is also probably quite unreadable for someone who is not familiar with mathematical logic. Formal means that the real formal is abbreviated by language (which is most mathematics from the hands of mathematicians who do not deal with real formal logic as a subject on itself). Right informal mathematics is when something is discussed in words for educational purposes or some parts of proofs are left out, which are proven somewhere else. Wrong informal is when something is stated without prove, which is not even formally true. This distinction will be discussed in the first section.

Moreover, considering notation, the *emphasized*, **bold** and 'quoted' texts are either new words/definitions or (general) important notions. Also, 'quoted' text can be the informal analogy, in words, of something that is being explained or the introduction of symbols, not necessarily text. I hope it will be clear in the text.

In addition, there is something to say about the (albeit subjective) beauty of mathematics. After continuity and some other concepts are clear, there is the possible division between the *continuous* and the *discrete*, and it is quite remarkable how these can overlap such as in algebraic topology.

Furthermore, this book does not claim to be a *complete* introduction. So try to consult other sources and taste other styles of doing mathematics. In general, I would encourage you to investigate some of the concepts on the internet or with other sources if they are not immediately clear. In mathematics there can be many different ways of proving the same theorem, so it is useful to see other viewpoints. Moreover, if there are any mistakes in this, which I have tried to minimise of course, do not panic, but think for yourself. Because that is ultimately what doing mathematics *is* about: to learn for yourself to know what formal mathematics is.

Finally, mathematics is not just isolated-cases-problem-solving, it also is about coming up with new problems/questions in trying to prove a general result. So you can ask yourself: do I want to be a generaliser and prove

theorems or do I want to do (long, tedious) calculations? But most of all, when doing mathematics, you should enjoy it!

Chapter 1

Proving theorems

In the first weeks of studying mathematics at university, you will most likely come across several exercises in which you need to give a *proof* of a certain *proposition*. *Proof theory* is a part of the field of mathematics, which studies the question: "What is a 'good/allowable' proof?". In the beginning of the 20th century, mathematicians established a formal proof theory. Their theory answers the raised question, according to today's standards. The main idea is that a proof is a line of reasoning establishing a certain proposition when it meets some or all 'proof-rules' and only 'axioms' and the explicitly mentioned presumptions are taken for granted (to be 'true').

We have the notations: '$\wedge$' means 'and', '$\vee$' means '(inclusive) or'. Now officially, these are the *real formal* symbols for 'and' and 'or', but they just represent (formal) *functions* associated with the idea of 'and' and 'or' with two inputs (propositions) and output 0 or 1: two inputs $(\alpha, \beta) \mapsto 0$ or 1, depending on the validity of α and β and the used comparing symbol. This function notation will come back later.

The following 'truth' table tells you what the 'truthness' is (0 or 1) of a combination of symbols when they are both true up to when they are both false etc. Beware that 0 or 1 doesn't actually have to mean something *semantically*[1]. It is only to make a distinction. This is what separates mathematics from mere linguistics.

The truth table:

Input		Output		
α	β	$\wedge$	$\vee$	$\rightarrow$
0	0	0	0	1
0	1	0	1	1
1	0	0	1	0
1	1	1	1	1

[1]Semantically means having to do with the 'meaning' of words.

But we can allow semantic discussions to make it (informally) clear:

$$\text{red is a color} \wedge \text{blue is a color} = \text{true},$$

$$\text{an apple is a pear} \vee \text{you wear red socks} = \text{unclear}.$$

Now the first example is (informally) obvious, but the second is (informally) unclear because if you wear red socks one of both statements is true, but if you do not wear red socks, clearly the combined 'or' statement is also false. However for $\vee$ to give true, also both input statements can be true, therefore it is the inclusive or. Moreover, I write (informally) because you cannot really easily define what a color is etc. Again mathematics is (trying) not (to be) philosophy or linguistics.

Sometimes we have a proposition which depends on some variables. We can write this as $\phi(x_1, x_2, ..., x_n)$. Example: $x_1 \times x_2 = x_3 \rightarrow x_2 = x_3 \times x_4$. Now this does not specify conditions on $x_1, x_2, \ldots, x_n$. For example $x^2 = 3$, does not 'mean' anything formally. We must have, for example,

$$\text{for all numbers, } x, \text{ in the set of positive integers: } x^2 = 3,$$

which is (informally) a *closed-formed* proposition, but it is false in its validity. We will come to this closed-formedness with (formal) quantifiers later.

The following 'prove-rules' are essential, and the most common prove-rules[2], which allows you to prove theorems: in the following α and β are 'propositions'

1. A simple rule is that if you have (by presumption) $\alpha \wedge \beta$, you may conclude both α and β. Formally: $\alpha \wedge \beta \Rightarrow \alpha$ and $\alpha \wedge \beta \Rightarrow \beta$. Where '$\Rightarrow$' means, 'I draw the right conclusion from the left'. Of course, this doesn't apply to '$\vee$'. Now the '$\Rightarrow$' and the '$\rightarrow$' are related, but the one is the informal notion of implies and the second the (real) formal.

2. Another simple one is that if you are given a proposition (by presumption) α, you can always use it. Formally: $\alpha \Rightarrow \alpha$.

3. If you know that α 'implies' β and α is true, then you may conclude β is true. This is called the 'modus ponens'. Formally $(\alpha \wedge (\alpha \rightarrow \beta)) \Rightarrow \beta$, where the brackets $(\ldots)$ are only to group things on the left, separated from the right.

4. If you assume α is true, and you can somehow derive β out of α, then you may conclude α implies β. Formally $([\alpha] \Rightarrow \beta) \Rightarrow (\alpha \rightarrow \beta)$, in

[2]One has the freedom to choose his/her own prove-rules, but the rules might not be 'allowable' by real formal logic. Now again 'real' formal logic (mathematical logic) is a subject on its own, however interesting, laying outside the scope of this book.

which [...] means that what is inside the brackets is an assumption. Be careful not to just use α as a proposition that is necessarily true. Now in a so-called *theory* in mathematical logic

5. If you have to prove α, you can assume that α is not true, and somehow derive a contradiction. If you find a 'general' contradiction which leads solely from you assumption that α is not true, you may conclude that α is true. Formally: $([\neg\alpha] \Rightarrow \bot) \Rightarrow \alpha$, in which $\bot$ means a proposition that is always false. Be careful as to how your assumption effects the contradiction, this is what I mean with 'general': it is a valid contradiction for all circumstances to be checked, not just in one example. This method is called 'Reductio ad absurdum (RAA)' or 'proof by contradiction'.

6. And, of course, rewriting a mathematical expression (using the '=' sign) can sometimes be enough to give a proof.

7. Sometimes you have to prove unicity of x in a proposition $\phi(x)$, which depends on x. A common way to proving this is assuming there is a y for which the proposition **also** holds, and deducing that $x = y$ by contradiction.

8. A different sort of technique to prove statements is 'induction'. This technique deals with expressions that take a natural number as input. It is a technique that proves that a proposition is true for all natural numbers. Induction starts with the assertion that an proposition, ϕ is true for the base case. Formally: $\phi(0)$. After this, the assumption $\phi(n)$ should imply $\phi(n+1)$. Examples will follow.

Something else you will come across often in mathematics is proving a 'if and only if' expression. Real formal notation: $\alpha \leftrightarrow \beta$. This is nothing else as the expression $\alpha \rightarrow \beta \wedge \beta \rightarrow \alpha$. So if you have to prove an 'if and only if' relation, you need to prove both ways. Easiest is to prove both ways separately. Writers use the symbol $\iff$ to imply an 'if and only if' relation.

Some of the following exercises belong to this section, but can be done later, when things become more clear.

Exercises

1. **Very Important** Think about this statement: $\alpha \to \beta \iff \neg\beta \to \neg\alpha$ in words (informally). Try proving it formally using the assumption $[\neg\beta]$ and $[\alpha]$ and using the expression $\alpha \to \beta$ as a true one (you always need to assume something in front of a '$\to$', do you see that?). Can you use RAA to change α to something desired? This is known as the 'contraposition'. Something very useful about this, is that if you have to prove $\alpha \to \beta$, you can also just prove $\neg\beta \to \neg\alpha$. Add it to your proving-techniques list!

2. The classic example for induction is $\phi(n) = \big(1 + 2 + 3 + \cdots + n = \frac{1}{2}n(n+1)\big)$. First check the base case $\phi(1)$. Then by assuming $\phi(n)$, rewrite $1 + 2 + \cdots + n + 1$, by using $\phi(n)$ and simplify the right side to the right side of $\phi(n+1)$.

3. Prove by induction that for all positive integers $n > 3$ that $3^n > n^3$.

4. Prove by induction that the number of diagonals of a regular polygon with n sides is $\frac{1}{2}n(n-3)$. Try to think of exemplary pictures of some small regular polygons first.

5. Prove by induction that the number of minimal moves to solve the Tower of Hanoi (search for it on the internet) for n disks is $2^n - 1$.

Chapter 2

Notation and definitions

2.1 Quantifiers

Now, quantifiers are formal objects, but again are (informal) analogies of the following:

1. $\exists_{x \in a} \beta$ means 'there exists a x in a certain set a such that β is true/valid'. (The $x \in a$ in subscript is only to make a clear distinction between $x \in a$ and β.)

2. $\forall_{x \in a} \beta$ means 'for all x in a certain set a, β is true/valid'.

Sometimes $\exists_{x \in a} \beta$ is rather subtle, because it also means something like: "if there is a x such that β, then, ... holds".

Now, a closed-formed proposition ϕ, depending on $x_1, x_2, \ldots, x_n$ (so $\phi(x_1, x_2, \ldots, x_n)$), is a proposition in which all variables are restricted by exactly either one of these quantifiers. So in our example in section 'Proving theorems', we had $\forall_{x \in \mathbf{N}}$. This will become more clear in the following section.

2.2 Sets

Sets are collections of items, informally. In fact, it can be said, jokingly, all mathematics is applied set theory (combined with the human capability to grasp certain concepts of course). Set theory was formally established at the same time as proof theory. The 'set' of axioms which tames the most used 'forms' of sets is called Zermelo-Fraenkel (**ZF**) set theory, which is just a list of axioms. Some of their axioms are derived from quite simple occurrences or thought experiments. One of them is 'Two sets are the same, when they have the same elements.', which is actually a very useful axioms. Another axiom prevents instances such as the 'Russell's paradox' (Google it I'd say!) Other axioms are used to define some notation and operations of set theory,

such as (in the following a, b and c are sets, x are elements, beware: elements can also be sets!):

1. A set in set-notation has two parts: the part left of the '|' and the part right of it: $\{... \mid ...\}$ where the brackets '{' are to 'close the set off'. The part left is to define in what set the elements of the set at hand are in (because this is required because of an axiom, for really big sets this is the 'Von Neumann Universe') and this is often left out. The part right is to give a condition for an element which it should obey. Otherwise, sets are denoted as the set with some explicit elements like this $\{a, b, c\}$, $\{1, 2, 4, 7\}$ or {apple, pear, cow}, this can only be done when dealing with small finite sets. Sometimes ':' can be used instead of '|'.

2. $\emptyset$ or $\{\ \}$ is the empty set. It is a set with no elements. It is guaranteed to exist by an axiom. This set doesn't have the two parts.

3. Sets in **ZF** don't care about double elements, thus $\{x, x, x\} = \{x\}$

4. $a \cup b = c$ is the union, c, of a and b. c contains all elements of both a and b. Formally: $c = \{x \mid x \in a \vee x \in b\}$. Again here, formally is not real formal, but is how it is written in a precise way (and hopefully understandable) using the notation we have come past. Now x is an element of the 'Von Neumann Universe', but this is not important when doing mathematics which is not always about set theory on its own. Moreover the notation $\bigcup_{x \in A} x$ means the union of all $x \in A$, example: Suppose $B = \{\{a, d\}, \{c\}, \{8\}, \{3, w\}\}$. Then $\bigcup_{x \in B} x = \{a, d, c, 8, 3, w\}$, much like putting all elements of elements of B into one set together.

5. $a \cap b = c$ is the intersection, c, of a and b. c contains all elements which are both in a **and** b (notice the subtlety). Formally:
$c = \{x \mid x \in a \wedge x \in b\}$. There is a similar notation for $\bigcap_{x \in A} x$, just like for the union.

6. $a \subseteq b$ or $a \subset b$, means a is a subset of b, i.e. all elements of a are contained in b: $\forall_{x \in a}\ x \in b$. '$\subseteq$' means '$a$ is equal to b' **or** 'a is a subset of b'. Because if $a = b$ then still a is **subset** of b, but not a *strict* subset. '$\subset$' is sometimes interpreted as a 'strict' subset, so that $a \neq b$, hence, the formalism with this notation is unclear. A means to make a distinction between strict subset and $\subset$ is using '$\subsetneq$' for strict and '$\subseteq$' for any subset instead.

7. $P(a) = c$ means: the *powerset* of a is c. Formally: $c = \{x \mid x \subseteq a\}$. Why is the empty set an element of $P(a)$? This is the set of all subsets of a.

8. $a \backslash b = c$ means: 'a without its elements which are also in b'. Formally: $c = \{x \mid x \in a \wedge x \notin b\}$

9. $a \times b = c$ is the Cartesian product of a and b, containing all **ordered** pairs (a_i, b_i) $\big(\neq (b_i, a_i)$ in general$\big)$, with $a_i \in a$ and $b_i \in b$. Formally: $c = \{(a_i, b_i) \mid a_i \in a \wedge b_i \in b\}$. The formal set notation of one ordered pair (x, y) is $\{\{x\}, \{x, y\}\}$. Remember, $\{x, y\}$ has no order.

10. In dealing with intervals of real numbers, a closed interval is $[a, b] = \{x \in \mathbf{R} \mid a \leq x \leq b\}$ and an open interval $(a, b) = \{x \in \mathbf{R} \mid a < x < b\}$.

A very **useful** technique to prove that two sets are equal, is to prove that both sets are subsets of each other. Formally: $(a \subset b) \wedge (b \subset a) \Rightarrow a = b$

Often used sets are: **N**, **Z**, **Q** and **R**, which are the sets of natural (positive) numbers, all whole numbers, rational numbers and the ominous real numbers respectively. They are also denoted as $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ and $\mathbb{R}$ respectively. Furthermore, $\mathbf{R}[x]$ and $P_n(\mathbf{R})$ are the set of all polynomials and the set of all polynomials up to a certain degree (degree is highest exponent of a polynomial) respectively, both with coefficients in **R**. Examples: $x^2 + \sqrt{2} \in \mathbf{R}[x]$ and $x^4 + 2x^3 \notin P_3(\mathbf{R})$. Generally, if R is a *ring* we let $R[x]$ denote the set of all polynomials with coefficients in the ring R. We will pass rings later.

Moreover, a nice construction is the *partition*. A partition of a set B, is a set A, for which holds:

1. $\forall_{x \in A}\ x \neq \emptyset$
2. $\bigcup_{x \in A}\ x = B$
3. $\forall_{x,y \in A}\ x \cap y \neq \emptyset \Rightarrow x = y$, or equivalently: $\forall_{x,y \in A}\ x \neq y \Rightarrow x \cap y = \emptyset$. This called: x and y are *disjoint*.

Now, a partition is really analogically a partitioning of all the set's its elements into disjoint subsets of the set.

Exercises

A, B, C and D are any sets.

1. Prove: $(A \cup B) \cup C = A \cup (B \cup C)$
2. Prove: $(A \cap B) \cap C = A \cap (B \cap C)$
3. Prove: $A \cap B = A \iff A \subseteq B$
4. Prove: $A \cup B = A \iff B \subseteq A$
5. Prove: $A \backslash B = A \iff A \cap B = \emptyset$
6. Prove: $A \backslash (B \cup C) = (A \backslash B) \cap (A \backslash C)$
7. Prove: $A \backslash (B \cap C) = (A \backslash B) \cup (A \backslash C)$
8. Prove: $P(A) \subseteq P(B) \iff A \subseteq B$
9. What does $\mathbf{R} \times \mathbf{R}$ resemble visually, thinking about the xy plane?
10. Prove: $(A \times B) \neq (B \times A)$ in general (which means: not always). Think about a counterexample.
11. Prove: $(A \cap B) \times (C \cap D) = (A \times C) \cap (B \times D)$
12. If $x \in A$ and $A \backslash B \subseteq C \cap D$, does $x \notin D \Rightarrow x \in B$?
13. Given $x \in C$ and $A \cap C \subseteq B$. Does this imply $x \notin A \backslash B$?

2.3 Relations

The way to formally "match" two elements, which are *in a relation* with each other, is easily done via ordered pairs. So if $x, y \in A$ then for some relation, you write xRy or $x \sim y$, whenever x has the specified relation with y. Some examples are $x \leq y \iff x \sim y$ or $x + y = \text{even} \iff xRy$. Beware to distinguish between the '=' for elements in sets and the '=' in the relation prescription. There are some types of relations which obey to certain common sense rules:

1. Equivalence relation $\forall_{x,y,z \in A}$:
   (a) Reflexivity: xRx
   (b) Symmetry: $xRy \Rightarrow yRx$
   (c) Transitivity: $xRy \wedge yRz \Rightarrow xRz$
2. Partial order $\forall_{x,y,z \in A}$:
   (a) Reflexivity
   (b) Antisymmetry: $xRy \wedge yRx \Rightarrow x = y$
   (c) Transitivity

If we have some equivalence relation we have an equivalence class for each $x \in A$, which is a set, denoted as '$[x]$', which contains all elements with whom x has a relation with. Formally: $[x] = \{a \in A \mid xRa\}$. The set of all equivalence classes together is called the *quotient set*, formally: $A/\sim := \{[x] \mid x \in A\}$, where ':=' means, '... is defined as ...'. Equivalence classes are often used in defining concepts in mathematics. Examples are the congruence classes, related to modular arithmetic. These congruence classes can be defined formally using group theory, which will come back later.

Another example are the equivalence classes induced by so-called *homotopy* in *topological spaces*, which has something to do with how a space is punctured by holes. This often referred to as in that a mug is topologically the same as a donut, because they both have one hole (and can be continuously deformed into each other).

Moreover, they appear in the concept of *orbits* in *group theory*, which is useful for counting for one thing.

Exercises

1. Give the quotient set of $x + y =$ even, with $x, y \in \mathbf{N}$.
2. **Very Important** Given an equivalence relation over the set A, prove that the quotient set is a partition of A. Reminder: set theory doesn't care about double elements.
3. Check that given the set $\mathbf{N}$ the relation less than or equal to '$\leq$' over this set, is a partial order on $\mathbf{N}$.
4. Check that given a set A, the relation subset or equal to '$\subseteq$' over $P(A)$, gives a partial order on $P(A)$.
5. **Very Important** Check that given a simple finite quotient set of some relation, such as in the examples/exercises above, you can choose a single *representative* element for every equivalence class in the quotient set. This is known as the finite Axiom of Choice, which is a controversial element of extended **ZF** theory when considering infinite sets. (Again, this is mathematical logic/axiomatic set theory.)

2.3.1 Defining some sets

Now we can actually define $\mathbf{N}$ using only sets. This can be made really formal but I will sketch some ideas here.

Let the symbol $0 := \{\}$. And $S(a)$ the successor function with operation $a \mapsto a \cup \{a\}$. Then $1 = 0 \cup \{0\}$ and $2 = 1 \cup \{1\}$. Then $n = n - 1 \cup \{n - 1\}$. This construction is thanks to Von Neumann. Another way is to define the natural numbers with axioms, such as the *Peano axioms*. It can be shown that Von Neumanns constructions satisfies the Peano axioms.

The integers $\mathbf{Z}$ can be defined using relations. Let $a, b \in \mathbf{N}$. We say $(a, b) \sim (c, d) \iff a + d = b + c$, using the definition of addition in $\mathbf{N}$. Now we can choose one representative for every equivalence class of this relation on $\mathbf{N}$ and call that n or $-n$. Here you can hold the analogy in mind that $(1, 3)$ is related to $1 - 3 = -2$ and $(0, 2) \sim (1, 3)$, so let positive integers n be $(n, 0)$ and negative $-n$ be $(0, n)$, for example. Because the representative is non-unique.

We can define the rational numbers $\mathbf{Q}$ with another equivalence relation. Let $a, b, c, d \in \mathbf{Z}$. We say $(a, b) \sim (c, d) \iff a \cdot d = b \cdot c$, with multiplication in $\mathbf{Z}$. Now we can again assign representatives for every equivalence class and call that some rational number.

2.4 Functions

Functions, f are 'objects' which relate two sets, X and Y, with each other via a certain rule. Formally: $f : X \to Y$. The '$\to$' has nothing to do with implications, in the case of their usage in functions. Some features of functions are:

1. X is the *domain* of the function f

2. Y is the *codomain* of the function f

3. Every function 'maps' ($\mapsto$) every single element in the domain to one single element in the codomain. One element in the domain cannot be mapped to two different elements in the codomain at the same time for functions.

4. The 'certain rule' is a prescription, which 'says' what the functions does with its elements in the domain. "To what element in Y do I, the function, have to 'map' this element in X?". It is very similar to giving a computer an algorithm to execute.

   Example: $f : \mathbf{N} \to \mathbf{N}$, thus $X = \mathbf{N}$ and $Y = \mathbf{N}$, with prescription: $x \mapsto x^2$. '$\mapsto$' means 'maps to'. Here clearly $x \in X$, and $x^2 \in Y$. According to our notation so far: $f(2) = 4$, in this case, right?

5. If $a \subseteq X$ then $f(a)$ is the *image* of f under a. Notice that $f(a) \subseteq Y$. The 'range' of f is $f(X)$. Here the subtlety of $\exists$ comes around, because the formal notation of the image is:

   $$f(a) = \{y \in Y \mid \exists_{x \in a} y = f(x)\}.$$

   In words this says something like: "If for some $x \in a$ and $f(x) = y$, then $y \in f(a)$". So the set, $f(a)$, loops, as to say, over the images of all elements in a and adds them to its set. Now $f(X) \subset Y$ of course.

6. If $b \subseteq Y$ then $f^{-1}(b)$ is the domain elements in X that map to b under f.

7. *Injectivity* means all function outputs are different from each other. Formally: $x_1 \neq x_2 \Rightarrow f(x_1) \neq f(x_2)$ or (the contrapo...) $f(x_1) = f(x_2) \Rightarrow x_1 = x_2$. Notice that $x_1 = x_2 \Rightarrow f(x_1) = f(x_2)$ makes not much sense because of feature 3.

8. *Surjectivity* means that the function maps 'onto' all elements of the codomain. Formally: $\forall_{y \in Y} \exists_{x \in X} f(x) = y$. Notice that surjectivity doesn't imply injectivity. If $f(X) = Y$ then f is surjective, right?

9. *Bijectivity* means that the function is both injective and surjective. What does this mean? Notice that every single element in the domain gets a different output, and every element in the codomain is mapped onto.

10. $h = g \circ f$ is the function that takes as input an element of the domain of f which gives an element in the codomain of f, and if the domain of g is the codomain of g, this g can give an output of this element. The mapping is like this (Z is the codomain of g): $h : X \to Y \to Z$ and $h(x) = g(f(x))$. Think about it as g AFTER f. We call this function-*composition*. We can check that function-composition is associative, i.e.: Let $f : A \to B, g : B \to C, h : C \to D$ then

$$h \circ (g \circ f) = (h \circ g) \circ f.$$

Because for some $x \in A$

$$h \circ (g \circ f)(x) = h \circ (g(f(x)) = h(g(f(x)) = (h \circ g) \circ f(x).$$

11. We define the set of all possible functions from $X \to Y$ as Y^X. This is interesting when X and Y are finite. Most of the time in talking about these sets, if someone writes 3^X, they mean for the $3 := \{0, 1, 2\}$. Example:
Suppose $B = \{a, b\}$, then $2^B = \{f_1 := (f_1(a) \mapsto 0, f_1(b) \mapsto 0), f_2 := (f_2(a) \mapsto 1, f_2(b) \mapsto 0), f_3 := (f_3(a) \mapsto 0, f_3(b) \mapsto 1), \dots\}$. So all possible mappings between the codomain and the domain are included.

12. A *linear* mapping $f : V \to W$ (just another term for a function) is such that $\forall_{u,v \in V}$:

$$f(u + v) = f(u) + f(v)$$

$$f(cu) = cf(u)$$

I have still left out the rigorous definition of the '+' and c, because that will become clear after you have become familiar with groups, rings and fields.

We say a set A is *finite* (in the number of its members) if there exists a bijection $f : A \to \{1, 2, \dots, n\}$ for some $n \in \mathbf{N}$. Now the number of its elements is commonly written as $|A|$ or $\#A$, which would be the number n.

Another interesting mapping is the permutation. If you have a function $\sigma : X \to X$, then any bijection σ is also a permutation of X. Example: $X = \{1, 2, 3, 4, 5\}$ then some permutation of the ordered quintet $(1, 2, 3, 4, 5)$, a change of order, could be: $(2, 5, 3, 1, 4)$. This is the same *structure* as the

bijection for which $\sigma(1) \mapsto 2$, $\sigma(2) \mapsto 5$, $\sigma(3) = 3$ etc. I.e. the structure contains the same information. You can denote a permutation as:

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 5 & 3 & 1 & 4 \end{pmatrix}$$

Or in general:

$$\sigma = \begin{pmatrix} x_1 & x_2 & x_3 & \dots \\ \sigma(x_1) & \sigma(x_2) & \sigma(x_3) & \dots \end{pmatrix}$$

Doing σ on some ordered list like $a = (1, 4, 3, 6, \dots)$ gives a new list:

$$\sigma(a) := \sigma = (\sigma(1), \sigma(4), \sigma(3), \sigma(6), \dots)$$

on which you can do a new permutation, which gives you:

$$\sigma\sigma(a) := (\sigma(\sigma(1)), \sigma(\sigma(4)), \sigma(\sigma(3)), \sigma(\sigma(6)), \dots)$$

etc. So you can repeat permutations after each other like multiplication. But why does order matter in contrast to multiplication? Actually, permutative multiplication is the composition of bijective functions.

A very useful proving technique due to bijections is the *pigeonhole principle*: there is no bijection possible between $\{1, 2, \dots, n\}$ and $\{1, 2, \dots, m\}$, with $n < m$.

Exercises

1. Prove the pigeonhole principle.
2. Give the formal set notation of $f^{-1}(b)$.
3. Prove that, if some function $f : X \to Y$ is bijective, if and only if f^{-1} is bijective. f^{-1} is the function that sends all elements in Y to their element in the domain: $f(x) = y \Rightarrow f^{-1}(y) = x$.
4. Show that if the same $f : X \to Y$ is bijective, then $f \circ f^{-1} = \mathrm{id}_Y$ and $f^{-1} \circ f = \mathrm{id}_X$. id_A is a 'identity' function that sends all elements to the element itself, so $\mathrm{id}_A : A \to A$ and $\forall_{a \in A} \mathrm{id}_A(a) = a$.
5. Suppose $f : A \to B, g : B \to C, h : C \to D$ are functions and $g \circ f$ and $h \circ g$ are bijective. Prove that there is a bijection $h \circ f$.
6. Prove that the function $T : \mathbf{R}[x] \to \mathbf{R}[x]$ with prescription: $T(f(x)) = f'(x)$ is not injective but is surjective. Where $f'(x)$ is of course the (single variable) derivative of $f(x)$.
7. Give a injection (a function that is injective) $f : \mathbf{N} \to \mathbf{Q}$.

8. Search for the 'pairing function' on (English) Wikipedia. Using this we can prove that there is a bijection $f : \mathbf{N} \to \mathbf{Q}$, which seems quite remarkable.

9. Give a bijection between all even numbers and $\mathbf{N}$.

10. Is this function bijective?
Let a be any set. $f : P(a) \to P(a)$, with $x \in P(a)$, $x \mapsto a \backslash x$

11. Show that if a is any set, then there is a bijection between $P(a)$ and 2^a (as in the set of all functions $f : a \to \{0, 1\}$).

12. Suppose you have: $\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$. What is $\sigma\sigma(1, 2, 3)$? This means you have found an *identity element* for σ. (Similar as with the identity function an identity element is an element that under operation with another element gives the original element, it is a 'neutral' element, as to say, we will come back to this in the next chapter.)

13. Check that permutation multiplication is associative.

2.5 Negation of propositions

Now that we have become familiar with notation in quantifiers, we can address *negating* a proposition. This is useful if you have to prove $A \to B$, but you think it is easier to prove $\neg B \to \neg A$ (the contraposition). Formally, if you have a proposition purely in formal notation the $\forall$ changes in $\exists$ and $\exists$ in $\forall$. And $>$ changes to $\leq$, etc. Think about this with a simple proposition, and negate it. The element inclusion condition in the subscript does not change. Example:

$$\forall_{x \in \mathbf{R}} \exists_{y \in \mathbf{Z}} \; xy > 2.$$

Whose negation is

$$\exists_{x \in \mathbf{R}} \forall_{y \in \mathbf{Z}} \; xy \leq 2.$$

Again, it does not matter whether one is true or not, it is about negating propositions. Another example:

$$\forall_{n \in \mathbf{N}} \exists_{x,y,z \in \mathbf{N}} \; x^n + y^n = z^n.$$

Whose negation is:

$$\exists_{n \in \mathbf{N}} \forall_{x,y,z \in \mathbf{N}} \; x^n + y^n \neq z^n.$$

Chapter 3

Topics of interest

3.1 Limits of real sequences

Sequences are important concepts in the foundation of real analysis, which is the foundations of Calculus, with which you are probably familiar. Sequences give rise to the definition of series. Series are important for functions such as $\sin x$ and e^x. Series actually define these functions formally. A real sequence is a simple function: $f : \mathbf{N} \to \mathbf{R}$, often denoted as $(a_n)_{n=0}^{\infty}$ or $\{a_n\}_{n\in\mathbb{N}}$ for just the elements of the sequence, with $f(n) = a_n = x$, with $n \in \mathbf{N}$ and $x \in \mathbf{R}$. But when are sequences interesting? When they converge to some number! This is the fundamental idea of all of Calculus. But what is convergence of a sequence? It can't just be: 'as n gets big, $f(n)$ is some number'. There are two very important definitions of convergence of sequences in $\mathbf{R}$.
First, lets introduce some terminology for sequences:

1. A sequence $(a_n)_{n=0}^{\infty}$ is said to be decreasing if for all n $a_{n+1} \leq a_n$ holds. Strict decreasing is when $a_{n+1} < a_n$ holds. There is an analogy for an increasing sequence.

2. Sequences are actually sets which contain all 'output' of the sequence itself as a subset of $\mathbf{R}$. A sequence $(a_n)_{n=0}^{\infty}$ is bounded if for all n $|a_n| \leq M$, with $M \in \mathbf{R}$.

3. It's very important to note that $\pm\infty \notin \mathbf{R}$. The algebraic operations you may use when we use $\bar{\mathbf{R}}$ ($\bar{\mathbf{R}} = \{-\infty\} \cup \{\mathbf{R}\} \cup \{+\infty\}$, this is 'extended' $\mathbf{R}$), are for $x \in \mathbf{R}$:

   (a) $x + (+\infty) = +\infty$ and $x - (+\infty) = -\infty$

   (b) $x(+\infty) = +\infty$ and $x(-\infty) = -\infty$

   (c) $x/(+\infty) = x/(-\infty) = 0$

4. A more general notation in dealing with subsets of $\mathbf{R}$ is the supremum or infimum (The definition can actually apply to all sorts of partially ordered sets). Definition:
   If $A \subset \mathbf{R}$ and $A \neq \mathbf{R}$, then an *upper bound* $M \in \mathbf{R}$ of A is such that: $\forall_{x \in A}\ x \leq M$. A *supremum* $L \in \mathbf{R}$ of A is such that for all upper bounds of A, lets say M, $L \leq M$. This is denoted as $\sup A = L$.
   The case for $\mathbf{R}$ actually stands on this definition because in $\mathbf{Q}$ there are subsets which do not have a supremum in $\mathbf{Q}$ it self. This gave Dedekinds motivation to define real numbers as later-called *Dedekind-cuts*:
   Suppose x is a real number, then the Dedekind-cut for this real number is: $x := \{q \in \mathbf{Q} \mid q < x\}$.

The first definition of convergence, is convergence towards some limit $l \in \mathbf{R}$:

$$\forall_{\epsilon>0}\exists_{n_0 \in \mathbf{N}}\forall_{n \geq n_0}|a_n - l| < \epsilon.$$

Shorter notation: $a_n \to l$ or $\lim_{n\to\infty} a_n = l$. We use the informal *metric* $|\ldots| : \mathbf{R} \to \mathbf{R}_{\geq 0}$ with $x \mapsto |x|$ ($|\ldots|$ the absolute value of x), which is dependent on the set you are working in[1]. For the real numbers it is the familiar absolute value, for complex numbers it is $\mathbf{z} \mapsto \sqrt{z\bar{z}}$.

The second one is Cauchy convergence[2]:

$$\forall_{\epsilon>0}\exists_{n_0 \in \mathbf{N}}\forall_{n,m \geq n_0}|a_n - a_m| < \epsilon.$$

A very important argument in proving things about converging sequences is the $\epsilon/2$-argument'. Example: If $a_n \to a$ and $b_n \to b$, then $a_n + b_n \to a + b$.
Proof:
For some n_0 $|a_n - a| < \epsilon/2$ for all $n \geq n_0$, because convergence holds for all $\epsilon > 0$. Also, by the same argument, $|b_n - b| < \epsilon/2$. Take the biggest n_0 of a_n and b_n. Using the triangle inequality: $|x + y| \leq |x| + |y|$, we get:

$$|a_n + b_n - a - b| = |a_n - a + b_n - b| \leq |a_n - a| + |b_n - b| < \epsilon/2 + \epsilon/2 = \epsilon$$

Thus: $a_n + b_n \to a + b$.

Another useful tool in talking about limits is the so-called 'sandwich' principle or 'squeeze' theorem. It is as follows: suppose you have three sequences, $(a_n)_{n=0}^{\infty}$, $(b_n)_{n=0}^{\infty}$ and $(c_n)_{n=0}^{\infty}$, and there exists a n_0 such that $a_n \leq b_n \leq c_n$ for all $n \geq n_0$ and $a_n \to l$ and $c_n \to l$ as $n \to \infty$, then $b_n \to l$.

Furthermore, so-called boundedness of a sequence is very important, especially in addition with the condition that the sequence is decreasing or increasing.

[1] A formal metric is a function $d : X \times X \to \mathbf{R}_{\geq 0}$ for any X, such that some conditions hold.

[2] These two definitions are only equivalent in *complete metric spaces*.

Theorem A (the monotone convergence theorem):
Suppose $(a_n)_{n=0}^{\infty}$ is a real sequence, is increasing and bounded, then $a_n \to \sup\{a_n \mid n \in \mathbf{Z}^+\}$ as $n \to \infty$. There is an analogy for decreasing and bounded sequences.

Series are sums of infinitely many numbers. We write $\sum_{n=0}^{\infty} a_n = a_0 + a_1 + a_2 + \ldots$, where n is the starting index, and a_n the n-th element in a sequence. A partial sum is simply a series that stops for a finite N, we write the sequence of partial sums $S_N = \sum_{n=0}^{N} a_n = a_0 + a_1 + \cdots + a_N$. We say a infinite series converges whenever the sequence of partial sums converge, as in the definition of convergence for plain real sequences.

Exercises

1. Prove theorem A. *Hint:* first we may show that for $M := \sup\{a_n \mid n \in \mathbb{N}\}$ that $\exists n \in \mathbb{N}$ such that $M - \epsilon \leq a_n \leq M$ for every $\epsilon > 0$

2. Conclude that the subset of $A \subset \mathbf{Q}$ and $A = \{q \in \mathbf{Q} \mid q < \sqrt{2} \text{ (or equivalent: } q^2 < 2)\}$, doesn't have a supremum in $\mathbf{Q}$, but does if $q \in \mathbf{R}$.

3. Give the definition of the infimum.

4. Prove that $\frac{1}{n} \to 0$ as $n \to \infty$ by giving an explicit condition for n_0 in terms of ϵ.

5. Prove the sandwich principle. *Hint:* we have for $\epsilon > 0$ that $|a_n - l| < \epsilon$ and $|c_n - l| < \epsilon$ for all n greater than some n_0 and so $l - \epsilon < a_n$ and $c_n < l + \epsilon$

6. Prove that if $a_n - b_n \to 0$ and $b_n - c_n \to 0$ as $n \to \infty$, then $a_n - c_n \to 0$. Thus, you can interpret this as a equivalence relation R over the set of all converging sequences:

$$(a_n)_{n=0}^{\infty} R (b_n)_{n=0}^{\infty} \iff a_n - b_n \to 0$$

7. Show that $\sqrt{n+1} - \sqrt{n} \to 0$ as $n \to \infty$.

8. Consider the following cases for some $(a_n)_{n=0}^{\infty}$, where $a_n \in \mathbf{Q}_{\geq 0}$ as $n \to \infty$:

   (a) $\frac{a_{n+1}}{a_n} \to 1$ (come up with two sequences that have this property and either converge or does not)

   (b) $\frac{a_{n+1}}{a_n} \to a$ with $|a| < 1$

   (c) $\frac{a_{n+1}}{a_n} \to a$ with $|a| > 1$

   What can you conclude about $\lim_{n\to\infty} a_n$?

9. What is $\lim_{n\to\infty} \sqrt[n]{a}$ with $a \in \mathbf{R}$?

10. Prove: if $n \in \mathbf{N}$ then $(1+\epsilon)^n > 1 + n\epsilon$ for all $\epsilon > 0$, using the binomial formula.

11. Using the previous exercise, prove: $n^{1/n} \to 1$ as $n \to \infty$.*Hint:* think about using $\epsilon = n^{-1/2}$

12. Prove that the series $\sum_{n=0}^{\infty} (\frac{1}{2})^n$ converges. Show what it converges to. (This is a *geometric* series.)

3.2 Graph theory

Graph theory is very different from analysis, but is a very animating subject, because it deals with everyday, common-sense objects: networks, which we will call graphs. Problems from the field of graph theory are: 'the four colour theorem', 'the travelling-salesman-problem', making optimal time-schedules, and deciding a minimum flow in a network to suffice some demand. So this has more applied mathematics, compared to the first few sections.

A (simple) *graph* is an ordered pair of two sets: the vertices/nodes V and the edges/lines E between these edges. A graph G is then denoted as $G = (V, E)$, with for example $V = \{1, 2, 3, 4, 5\}$ and
$E = \{\{1, 2\}, \{3, 1\}, \{5, 4\}, \{3, 4\}, \{2, 3\}, \{3, 5\}\}$, represented visually by:

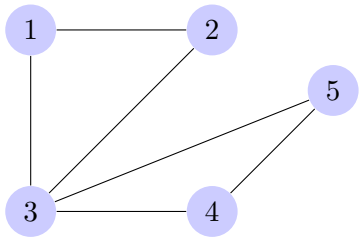

Figure 3.1: Example graph

The following is an enumeration of graph theoretical definitions. The **degree** of a vertex is the number of edges it is present in, visually: the number of lines that go out of the vertex. A k**-regular** graph is a graph with for every vertex the same degree k. A **complete** graph is a graph such that all vertices are connected, we write for the complete graph on n points: K_n. A **complete bipartite** graph is a graph $G = (V, E)$, such that there exists a partitioning of V in V_n and V_m, such that all points in V_n are *only* connected to *all* points in V_m, and thus the points of V_m only to V_n. For a complete bipartite graph with two such partitioning subsets with number of elements n, m we write $K_{m,n}$. An example is in figure 3.2. The **complement** of a graph $G = (V, E)$ is the graph $\bar{G} = (V, E')$ with E' the set of all edges that are not in G but are in the complete graph of G. A bipartite graph is a graph such that there are two subsets of V, say V_1 and V_2, such that

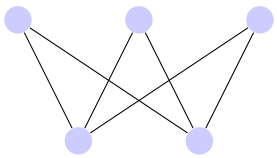

Figure 3.2: $K_{3,2}$

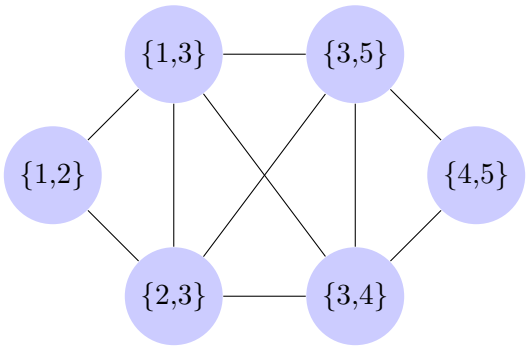

Figure 3.3: Example linegraph

$V_1 \cap V_2 = \emptyset$, and $\forall_{v \in V_1 \wedge w \in V_2} \{v, w\} \in E$ and $\forall_{v_i, v_j \in V_1} \{v_i, v_j\} \notin E$, and similar for V_2. So a bipartite graph is a partitioning on the vertices, such that they both 'connect'. A **walk** is a sequence of vertices $(v_0, v_1, v_2, \ldots, v_k)$, such that $\forall_{i \in N} \{v_i, v_{i-1}\} \in E$, with $N = \{1, 2, \ldots, k\}$. A **path** W is a walk such that $\forall_{v_i, v_j \in W} v_i \neq v_j$. Where the set W is sloppily used as the set of all vertices that are contained in the walk W. A **closed** walk, is a walk such that $v_0 = v_k$. A **circuit** is a closed walk, such that all vertices are different from each other. A graph is **connected** if $\forall_{v,w \in V}$ there exists a path between v and w. A **tree** is a connected graph without a circuit. Two graphs $G = (V, E)$ and $G' = (V', E')$ are **isomorphic** if there is a bijective mapping $f : V \to V'$, such that $\forall_{\{v,w\} \in E} \{f(u), f(w)\} \in E'$. The **linegraph** of a graph $G = (V, E)$ is a graph $L(G) = (V', E')$, such that $V' := E$ and

$$\forall_{e,f \in E} \; e \text{ and } f \text{ have a point in } G \text{ in common } \iff \{e, f\} \in E'.$$

See figure 3.3 for an example of the linegraph of the graph in figure 3.1. A **vertex-colouring** of a graph $G = (V, E)$ is a mapping $f : V \to C$, with $C = \{0, 1, 2, \ldots, k\}$ and $k-1$ the number of colours, such that $\forall_{u,v \in V} \{u, v\} \in E \to f(u) \neq f(v)$. The number $\chi(G)$ is the minimal colour number of G, which means that there exists a colouring with a possible minimum number of colours. The number $\chi'(G)$ is the **line-colouring** number, analogous to vertex colouring. A line-coloring is also a mapping $f : E \to C$ with $\forall_{u,v \in E} u \cap v \neq \emptyset \to f(u) \neq f(v)$. A **planar** graph is a graph without edges crossing each other, but how can we know if a graph is planar? Well, there is a theorem about this, which we will deal with in the exercises. The **facets** of a planar graph are the regions a planar graph is surrounding, the outside region included. So, for a triangle (K_3) the number of facets f is 2, the one it is surrounding and the outside. For K_4 it is 4.

Using induction we are going to prove Euler's characteristic for polyhedrons: $|V| + f = |E| + 2$. We use induction on $n + m$ of graph $G = (V, E)$, with $n = |V|$ and $m = |E|$. Check the base case $n + m = 1$. Now if $n + m > 1$,

we have two possibilities:

Case 1: G contains no circuits. Thus we have a tree. Then there is a point with degree 1 (see previous exercise). If we delete this point and its connected line, we have $n-1$ points and $m-1$ lines and the same number of facets. According to our induction hypothesis, we have $(n-1)+f = (m-1)+2$, implying $n+f = m+2$.

Case 2: G contains a circuit C, pick a line on this C, $\{u, v\}$. If we delete this line from G, the graph will still be connected. The n will stay n and we have $m-1$. The number of facets will be $f-1$. According to our induction hypothesis we have $n+(f-1) = (m-1)+2$, implying $n+f = m+2$.

A **directed** graph is an ordered pair of two sets: the arrows (directed) and the vertices. So we have $D = (V, A)$, where A is a set of ordered pairs $(v, u) \in A, v, u \in V$, because direction matters. A **flow** from the vertex s (source) to t (terminal) is a mapping $f : A \to \mathbf{R}^+$, so that

$$\forall_{v \in V \setminus \{s,t\}} \sum_{a \in \delta^-(v)} f(a) = \sum_{a \in \delta^+(v)} f(a).$$

The problem of finding a minimum flow in a given network with given capacities is a famous optimization problem. An understandable algorithm for this is the Ford-Fulkerson algorithm.

Exercises

1. Why can't there be a 3-regular graph on 5 vertices?
2. Given a graph $G = (V, E)$ with n vertices and degrees $d_1, d_2, \ldots, d_n$, how many edges does this graph have?
3. A **triangle** is a circuit of length 3. Show that if a graph $G = (V, E)$ has six vertices either G or $\bar{G}$ contains a triangle.
4. Given a complete graph $G = (V, E)$ with n vertices. How many unique paths are there between two points in this graph?
5. Show that a 3-regular Hamiltonian graph has the property: $\chi'(G) = 3$.
6. Show that if G is a tree, then G has at least one point with degree 1.

3.3 Groups, rings and fields

Groups, rings and fields are very interesting objects in mathematics. Formally they are sets under some operations. In the case of a 'group', it is under only one operation, and for a field and a ring, it's two operations. Groups are less complex than rings, in terms of the number of rules they have to obey. And rings are in the same way less complex than fields. If you know simple algebra you actually already used fields and groups, because $\mathbf{R}$ under addition and multiplication is a field.

These are the rules for a group over a given set G and a given operation '$*$' between two elements, beware '$*$' is just notation, but most of the times '$*$' is a familiar operation such as addition ('$+$') or multiplication:

1. Closure: $\forall_{x,y\in G}\ x*y \in G$. The notation for an operation could also be $*(x,y)$ or just xy, which returns a value in G, so it is just a function $* : G \times G \to G$.

2. Associativity: $\forall_{x,y,z\in G}\ (x*y)*z = x*(y*z)$.

3. Identity element: $\exists_{e\in G}\forall_{x\in G}\ e*x = x*e = x$. So for every element in G there is one neutral element in G which preserves the element.

4. Inverse element: $\forall_{x\in G}\exists_{x^{-1}\in G}\ x*x^{-1} = x^{-1}*x = e$

In fact, closure is a rather superfluous condition, because the operation demands $* : G \times G \to G$, which can be said to be a sort of 'closed' operation already. Furthermore, we say a subset H of a group G is a *subgroup* of G ($H < G$) if $\forall_{h_1,h_2\in H}\ h_1h_2 \in H$ and $\forall_{h\in H}\ h^{-1} \in H$. Moreover, if $x*y = y*x$ for every $x, y \in G$ for a group G then G is an *abelian* group, i.e. the group operation is *commutative*. A mapping $\phi : A \to B$ (function) between two groups A and B with operations $\circ$ and $*$ respectively is a *(group)homomorphism* if

$$\forall_{x,y\in A}\ \phi(x \circ y) = \phi(x) * \phi(y).$$

And if ϕ is a bijective mapping, then ϕ is a *(group)isomorphism*. We say the *kernel* of a group homomorphism $f : G_1 \to G_2$ is the set

$$\{g_1 \in G_1 \mid f(g_1) = e_{G_2}\}.$$

Where e_{G_2} is the identity element of G_2. We write $\ker(f)$ for the kernel of f.

The rules of a group also account for a field, but a field is more subtle then a group. So the rules of a field over a set F with operations '$+$' and '$\times$', which do not need to be actual addition and multiplication ('$\times$' is often left out as just xy, instead of $x \times y$, I will do this too), are $\forall_{x,y,z\in F}$:

1. Closure under + and $\times$
2. Associativity of both operations: $(x+y)+z = x+(y+z)$ and $(xy)z = x(yz)$.
3. Commutativity of both operations: $x + y = y + x$ and $xy = yx$.
4. Identity elements for both operations 0 (not necessarily $0 \in \mathbf{R}$) for + and 1 (not necessarily $1 \in \mathbf{R}$) for $\times$: $x + 0 = x$ and $x1 = x$.
5. Inverses for +: $\exists_{-x \in F}\ x + (-x) = 0$. Beware not to just simplify this expression to $x - x$.
6. Inverses for $\times$ except for 0: $\exists_{x^{-1} \in F}\ xx^{-1} = 1$. Beware not to simplify this expression to x/x.
7. Distributivity for + and $\times$: $x(y + z) = xy + xz$

For some notable groups we write the following:

1. The integers under addition $\mathbf{Z}^+$.
2. The rational numbers under multiplication $\mathbf{Q}^\times$, so this $\mathbf{Q}\backslash\{0\}$, because 0 does not have a multiplicative inverse.
3. The real numbers under addition and multiplication: $\mathbf{R}^+$ and $\mathbf{R}^\times$.
4. The complex numbers under addition and multiplication: $\mathbf{C}^+$ and $\mathbf{C}^\times$.

Now *rings* are sets equipped with two operations $(+, \times)$ just like fields. But now, only a ring R is an abelian group under + but not necessarily under $\times$. For $\times$ the following must hold for R to be a ring:

1. Left and right distributivity: $x(y+z) = xy+xz$ and $(y+z)x = yx+zx$
2. There is a $1 \in R$: $1x = x$ for all $x \in R$
3. Associativity: $x(yz) = (xy)z$ for all $x, y, z \in R$.

Rings are interesting in the sense that they allow some sort of an abstraction of the concept of prime/irreducible elements to be defined for any ring.

Exercises

1. Prove that the identity element of a group is unique. Prove that the inverse of an element is unique.

2. Prove that the subgroup of a group is a group.

3. We have a group homomorphism $f : G_1 \to G_2$. Prove that $\ker(f) \subseteq G_1$ is a subgroup of G_1.

4. Check that $\mathbf{Z}$ is a group under addition as you know it.

5. Check that $\mathbf{C}$ is a field under complex addition and multiplication.

6. Prove that if G is a group and $H \subset G$ is a finite subset, such that $\forall_{a,b \in H}\ ab \in H$ and $H \neq \emptyset$, then H is a group.

7. Suppose we have $\phi : \mathbf{R}^+ \to \mathbf{R}_{>0}$ with $x \mapsto e^x$, with $\mathbf{R}_{>0}$ the real numbers greater than zero under multiplication. Prove that ϕ is a grouphomomorphism.

8. Given some square, as you know it, check that by doing any reflection over a symmetry axis or leaving it (the identity element), gives you something which is still a square. Conclude that the square under any reflection is a group. In general, for any regular polyhedron with n sides this is the *dihedral* group D_n.

9. Let a be some set and $S = a^a \cap \{\text{functions } f : a \to a \text{ that are bijective}\}$. Check that $\forall_{x \in S}\ x$ is some permutation σ of a. Check that S is a group under the product $* : S \to S$, $\sigma_1, \sigma_2 \in S\ \sigma_2 * \sigma_1 = \sigma_2 \circ \sigma_1 = \sigma_2(\sigma_1)$. (as in the definition of permutations), for some example of a (not $a = \emptyset$ or $a = \{1\}$ or some other trivial easy example).

10. Check that $\mathbf{Z}[x]$, where these are the polynomials with **integers** coefficients, is a ring. (Any $a_0 + a_1 x + \ldots a_n x^n \in \mathbf{Z}[x]$ if $a_i \in \mathbf{Z}$ for all $0 \leq i \leq n$.)

3.3.1 Normal subgroups and quotient groups

We say subgroup N of group G is a *normal* subgroup if: $\forall_{n\in N}\forall_{g\in G}\ gng^{-1} \in N$. We write $N \triangleleft G$. A (left) *coset* of a subgroup in a group (of which it is a subgroup) is the **set** for some $g \in G$:

$$gH = \{gh \mid h \in H\}.$$

Exercise: check that the relation $\sim$ on the set of a group G for some subgroup H: $a \sim b \iff \exists_{g\in G}\ a, b \in gH$, is an equivalence relation. With this definition we can define *quotient* groups. First notice that the partition of the equivalence relation gives a representative element per (left) coset of H is G, together in the set S, such that $\bigcup_{s\in S}\ sH = G$. For the set of all (left) cosets of H in G we write $G/H = \{gH \mid g \in G\}$. If we have a representation set S for G/H we sometimes write $\overline{s}$ for sH. In general we can write $\overline{a}$ for aN. Exercise: N is a normal subgroup of $G \iff \forall_{a\in G}\ aN = Na$. Let N be a normal subgroup of G, then we define an operation on G/N as follows: $a, b \in G\ (aN)(bN) \equiv abN$. We have to check that it does not depend upon our choice of representative (well-definedness): if $aN = xN$ and $bN = yN$ (remember leftcosets are a partition of G), then

$$(xN)(yN) = xyN = x(yN) = x(bN) = x(Nb) = (xN)b = (aN)b = a(Nb) = abN.$$

Where we used that N is a normal subgroup of G. Exercise: check that G/N is group under this (well-defined) operation. Exercise: check that the *canonical* surjection $\phi : G \to G/N$ for $g \mapsto gN$ is a group homomorphism.

Now we can prove the fundamental theorem of homomorphisms together:

Theorem. Let G, H be groups and $f : G \to H$ a group homomorphism, let N be a normal subgroup of G and $\phi : G \to G/N$ the canonical homomorphism. If $N \subset \ker(f)$ then there exists a unique homomorphism $h : G/N \to H$ such that $f = h \circ \phi$.

Proof: we write $\overline{a} = aN = \phi(a)$ for $a \in G_1$. Define $g : G_1/N \to G_2$ by $g(\overline{a}) = f(a)$. Exercise: check that our choice does not depend upon our representative, i.e.: $\overline{a_1} = \overline{a_2} \Rightarrow f(a_1) = f(a_2)$. Show that g is group homomorphism. We now have $g \circ \phi(a) = g(\overline{a}) = f(a)$ for all $a \in G_1$. So $g \circ \phi = f$. Uniqueness: if $g' : G_1/N \to G_2$ and $g' \circ \phi = f$, then $g'(\overline{a}) = g'(\phi(a)) = g' \circ \phi(a) = f(a) = g(\overline{a})$. So $g' = g$.

Using the quotient groups we can formally define $\mathbf{Z}/n\mathbf{Z}$, which is modulo arithmetic with respect to $n \in \mathbf{Z}$. We can take subgroup $H = n\mathbf{Z} = \{nx \mid x \in \mathbf{Z}\}$ and look at the cosets $g + H$ with $g \in \mathbf{Z}$.

Exercises

1. Suppose H is a subgroup of the group G. Check that $aH = bH \Rightarrow ab^{-1} \in H$. And $aH = bH \vee aH \cap bH = \emptyset$.

3.3.2 Fields

R is a well known field, but contains pretty much elements. It is interesting that there exist fields which rely on only finite elements. An example:
A field modulo to some prime number p is written as $\mathbb{F}_p$. Its set contains all elements modulo to the prime. Example:
We have $\mathbb{F}_5 = \{\bar{0}, \bar{1}, \bar{2}, \bar{3}, \bar{4}\}$, with $\bar{a}$ the remainder modulo 5. Now we can write a multiplication and addition table and show that it is in fact a field:

$+$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$
$\bar{0}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$
$\bar{1}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$	$\bar{0}$
$\bar{2}$	$\bar{2}$	$\bar{3}$	$\bar{4}$	$\bar{0}$	$\bar{1}$
$\bar{3}$	$\bar{3}$	$\bar{4}$	$\bar{0}$	$\bar{1}$	$\bar{2}$
$\bar{4}$	$\bar{4}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$

The multiplication table is left to the reader as an exercise.

Now, finite fields are interesting for that they can be helpful in saying things about non-finite rings in the subject of *polynomial factorisation*. Moreover, they find application in cryptography. Furthermore, general fields form the elements of vectors of linear spaces.

A *vector space* or *linear space* over a field F is a set V with two operations: vector addition $+ : V \to V$ and scalar multiplication $\cdot : F \to V$ such that $\forall u, v, w \in V$ and all $a, b \in F$ we have:

1. (Associativity of vector addition) $u + (v + w) = (u + v) + w$
2. (Commutativity of vector addition) $u + v = v + u$
3. (Identity of vector addition) there is a $\vec{0} \in V$ (zero vector) such that $v + \vec{0} = v$
4. (Inverse of vector addition) there is a $-v \in V$ such that $v + (-v) = \vec{0}$.
5. (Identity of scalar multiplication) $1v = v$ where $1 \in F$ such that 1 is the multiplicative identity in F
6. (Compatibility) $a(bv) = (ab)v$
7. (Distributivity of scalar multiplication with respect to vector addition) $a(u + v) = au + av$
8. (Distributivity of scalar sums) $(a + b)v = av + bv$

Examples of common linear spaces are $\mathbf{R}^n$ and $\mathbf{C}^n$. Linear spaces are studied in the subject of linear algebra.

Exercises

1. Prove that every element of a finite field appears only once in every row and column in the multiplication and addition tables.

2. Prove that in a field $\mathbb{F}$ for $x, y, z \in \mathbb{F}$: $x + y = x + z \Rightarrow y = z$.

3.4 Topology of the real numbers and metrics

The topology of the real numbers leads us to define imporant concepts you have probably already heard of, such as *continuity* of a function or the *limit* of a function. This all leads to a nice theorem: every power series $\sum_{n=0}^{\infty} a_n z^n$ is continuous inside of its radius of convergence.

For a subset $A \subset \mathbf{R}$ we say, $x \in \mathbf{R}$ is a closure point of A if $\forall_{\epsilon>0}\exists_{y\in A}|x-y| < \epsilon$. Here the definition of the operator $|\ldots| : \mathbf{R}\times\mathbf{R} \to \mathbf{R}_{\geq 0}$ is important. For $\mathbf{R}$ we use

$$(x, y) \mapsto \begin{cases} x - y & \text{if } x \geq y \\ y - x & \text{if } y > x. \end{cases}$$

We write $\overline{A}$ for the set of all closure points of A. Now, if $a \in \overline{A}$, there exists a sequence $\{a_n\}_{n\in\mathbb{N}} \subseteq A$ such that $a_n \to a$ as $n \to \infty$.

We name the *epsilon neighbourhood* $a \in \mathbb{R}$ the set:

$$N_\epsilon = \{x \in \mathbb{R} : |x - a| < \epsilon\}.$$

We say $A \subseteq \mathbb{R}$ is *open* if $\forall_{a\in A}\ N_\epsilon(a) \subseteq A$.

Now we can say that a function $f : A \to B$ with $A, B \subseteq \mathbf{R}$ is *continuous* at $a \in A$ if

$$\forall_{\epsilon>0}\exists_{\delta>0}\forall_{x\in X}\ |x - a| < \delta \Rightarrow |f(x) - f(a)| < \epsilon.$$

Now, the $\forall_{x\in X}\ |x - a|$ can be interchanged with $\forall_{x\in N_\delta(a)}$.

Again, a *metric* on a set X is a function $d : X \times X \to \mathbf{R}_{\geq 0}$ such that the following hold:

1. $d(x, y) = 0 \iff x = y$

2. $d(x, y) = d(y, x)$

3. (Triangle Inequality) $d(x, z) \leq d(x, y) + d(y, z)$

From analysis we learn that the notions of convergence can all be stated with any well-defined metric. You may see this by looking at the proofs and the use of the triangle inequality in the case in $\mathbf{R}$.

(Provisional) Ending words

Now, the ending of the last section might seem a little abrupt, but this has a reason. If this booklet will receive much (positive) response, I will consider rewriting parts or writing more about other topics of interest. So, for now, I hope you enjoyed the first version of this booklet.