LAST
Logic
and
Axiomatic Set Theory
Lecture Notes in progress
2006, 1st quarter
## Contents

*Preface*

1. **About mathematics**
2. **Basic constructs of naive set theory**
3. **The infinite**
   - 3.1 The child
   - 3.2 Comprehension
   - 3.3 The hotel
   - 3.4 The ballroom
   - 3.5 Properties of equipotence
   - 3.6 Counting to infinity
   - 3.7 Ordinal and cardinal numbers
   - 3.8 Exercises
   - 3.9 Some further exercises
4. **Axiomatic set theory**
   - 4.1 Why axiomatize?
   - 4.2 ZFC—a survey
   - 4.3 Elementary constructions
   - 4.4 Ordinal numbers
   - 4.5 Induction and recursion
   - 4.6 Ordinal arithmetic
   - 4.7 Natural numbers and the set of natural numbers
5. **AC and some applications**
   - 5.1 AC - the Axiom of Choice
   - 5.2 Some direct applications
   - 5.3 Zermelo’s well ordering principle and two immediate applications

---

---
5.4  The Baire hierarchy 60
5.5  Zorn’s lemma 61
5.6  Some applications of Zorn’s lemma 63
5.7  Equivalences in ZF 65
5.8  Inadequacy of sequences for general topology 65
5.9  Generalized sequences, nets 67
Preface

These notes are designed to serve the needs of students of mathematics at the undergraduate level who take an interest in the basics as seen not just from a purely technical point of view but also with an interest in the meta aspects, the philosophical perspective.

Originally I intended that the notes should cover the basics of mathematical logic as well as set theory, aiming at a level corresponding to the requirements of the course “LAM” (which I refer to as “LAST”) under the new curriculum from 2004 at the University of Copenhagen. However, I chose, after all, to point to a standard textbook for the logic part of the course, viz. to Ebbinghaus, Flum and Thomas: “Mathematical Logic” (Springer, 1994). For the set theoretic part, my notes in Danish “Matematik Y, Introduktion til abstrakt matematik” from 2002 would have done nicely. Language requirements have led me to embark on a translation and revision of these notes - the result you see here. The student who wants to see an independent source is recommended to consult Hrbacek and Jech: “Introduction to Set Theory” (1.st ed. Dekker 1984).

Warning: I write these notes on a running basis. Hence the text as well as this preface may change as we go along. The homepage of LAST will guide the student and prevent, so I hope, confusion. The students of LAST are encouraged to assist me in the attempt – bound to fail, though – to reach perfection. But we may get closer. For the benefit of present and coming students and instructors.

-FT, August 27, 2006
1
About mathematics

Mathematics, what is it?

Mathematics is the science of structure and relations among structures which can be studied abstractly and lead to precise results through reasoning.

The emphasis on abstraction means that, essentially, mathematics is not dependant on the real world, on our senses, but goes on in our mind. A mind science. But, of course, mathematics is exercised by human beings placed in the real world with all its challenges which influence mathematicians as well as everybody else. This hardly points to abstraction as a goal in itself. And it is not. Rather, experience tells us that insistence on the possibility of abstraction is fruitful. Thus the word “can” in the above characterization of mathematics is important. No one says that mathematicians must always work in an entirely abstract universe – they don’t – but unless what they work with can be put into an abstract framework, it is not mathematics. The focus on abstraction in our definition excludes it being one of the physical sciences, but some human endeavors such as philosophy and theology may also fit this abstract focus. However, the last part of the definition with its reference to reasoning and precise results, interpreted as results everybody can agree upon, rules out that possibility.

Mathematics, why bother?

There are two reasons to study mathematics: As a tool to gain insight into other areas and as knowledge in its own right.
About mathematics

So one side of mathematics is as a (humble?) servant, assisting the other sciences. Many have wondered why the abstract mathematics is of any assistance, even essential for other sciences which strive to understand what can be learned through our senses. Well known is Nobel Laureate Eugene Wigner’s discussion paper: “The unreasonable Effectiveness of Mathematics in the Natural Sciences” (Wigner: 1902-1995). But perhaps it is not so strange. By letting the particular retreat into the background one achieves that general structure, the really essential features, sticks out and can be identified. In other words, the essence is given a chance to emerge through the mist of confusing details.

Mathematics, what do you need?

To exercise mathematics you need three elements: Methods for reasoning, something to reason about and motivation.

The methods for reasoning are covered by mathematical logic and the “something to reason about” you can think of as a chest, which we shall refer to as the treasure box† containing every object a mathematician could need. So the chest should contain all the numbers, but also objects as the sine-function, the Euclidean 3-dimensional space, groups, topological spaces, measure spaces, graphs etc., etc. In this view we see mathematics as a children’s game: Logic gives the rules of the game and the treasure box contains all objects which may enter into the game.

According to a very puristic view of mathematics, all you need is logic and objects for your reasoning, the treasure box. But really, this is a degrading of mathematics to a sterile game in an abstract universe devoid of any connection to the real world. Thus, when above we speak of motivation as the third element which enters into mathematics, this refers to the connection to the outer world and is represented by our intentions, our aspirations, our goals as human beings.

We emphasize the importance of the third dimension without which mathematics would make no sense. However, the two first elements form the foundation which makes mathematics possible and it is these basic elements which constitute the subject matter of the present notes. Experience has it that foundational issues have a great attraction on many – students and researchers alike. The philosophical interest largely lies in the fact that the foundations of mathematics enables man to loosen the dependence on the concrete, on the real world and allows man to create abstract universes without being restricted by the possibilities

† danish: @kramkisten
in the physical world. To make sense of the abstract universes one can build is another matter as already indicated.

The view of mathematics with only two foundational elements matured around 1900 with researchers like Cantor (1845-1918), Frege (1848-1925), Zermelo (1871-1953) and Russell (1872-1970). The importance of logic has long been recognized and dates back especially to classical Greece with Aristotelles (384-322 B.C.) as the key figure. But the second element, in our terminology the treasure box, is only some 100 years old. The idea is one of unification — once and for all to fix the objects mathematicians can play with. The notion of a set is in the foreground.

Adding also the notions of identity and membership (“$\in$”) as undefined notions and tying the notions together by carefully selected axioms, the idea is that then everything else in mathematics is to follow.

Before the “everything else” can follow one then has to agree on certain rules and axioms, a non-trivial task. For one thing, one has to work abstractly without recourse to physical models. Fact is that no one has ever seen the treasure box to contain all the sets we are to work with. It exists only in our mind. So all the wonderful precise facts of mathematics — $2 + 2 = 4$ (see exercise xx.xx), Pythagoras theorem and much, much more — rely on belief. Mathematics as religious belief! A degradation some would say. It is not. Instead, it points once and for all to a basis one has to believe in. Then everything else is indisputable. And what we have to believe in is judged to be natural by mathematicians. It is also comforting to know that, despite more than 100 years of research in the foundations of mathematics, no inconsistencies have yet been found.

In the nature of things, and despite the great freedom and flexibility provided by the foundations, there are limits to our power of reasoning, and to the ability to decide on every conceivable question. A bit disturbing. In Chapter xx we embark on a more critical study of these aspects, most prominently being pointed out to the mathematical community by Gödel (1906-1978).

We end this introductory chapter by illuminating a question often asked regarding greatness in science.

*Mathematics, what is the thrill of it?*

| Greatest in mathematics, as in science in general, is discovery — yet, foundations of mathematics is human invention. |

To emphasize our view let us recall what Archimedes (284-212 B.C.) wrote to his colleagues in Alexandria when referring to his discovery
that the area of the surface of the solid ball is 4 times the area of a cross section (at equator): “Ever since the dawn of times has the surface of a ball been 4 times a cross section, but I am the first to know it”† Not 4.021 times a cross section, no exactly 4 in our (Archimedes’!) abstract universe, an ideal universe, reflecting absolute truths about the real world. Archimedes’ proud announcement bears witness of the great moment it was for him to discover one of nature’s truths. Also, studying Archimedes’ path to the discovery fills us with admiration. In fact, his method was not paralleled until some 1800 years later with the invention (not discovery!) of the infinitesimal calculus.

As expressed above, discovery is the greatest. But the whole subject matter of these notes is devoted to those human constructions of the mind which constitute the foundation of mathematics. In the spirit of Wigner you could wonder about the “unreasonable” success of the inventions behind the foundation of mathematics as a tool to discover the truths about the real world. But, again, the inventive element – freeing itself from the limitations and distractions of the real world – is perhaps by necessity what is needed to create just the right tools to enable discovery.

One may argue that absolute truth is something we cannot know about. For one thing, the real world is only something we can get at indirectly - if it exists at all! In this view, only human inventions result from scientific efforts. We refuse to take such a puristic view. Examining examples of scientific development, the quality as invention or discovery often sticks out and can be judged by – admittedly subjective – grounds. We maintain that “You know it when you see it”.

Recognizing the foundations as merely human invention it follows that there is no unique path to the foundational tools. Our choice depends on the historical development and present day views of a majority of mathematicians. Perhaps some other day the foundation will be replaced by another foundation, generally agreed to be superior. Perhaps because it is easier to grasp, perhaps because it provides better tools to reach out to discoveries not easily accessible by the previous foundation. Or perhaps because it is more beautiful, aesthetically appealing to those who lead the way.

† slightly amended citation after xxxx.
Exercises

1.1 Clearly there are other ways to look at mathematics than here presented. Search the net for other definitions, compare with ours and discuss, based also on your own views! (For instance, you may want to find out who said that “Mathematics may be defined as the subject in which we never know what we are talking about, nor whether what we are saying is true.”)

1.2 Search the net using “Jaffe-Quinn” as the entrance term. You will find a recent debate among top-class mathematicians about the nature of present day mathematics. Aspects not touched upon above, e.g. reflecting the role of the computer and computation in mathematics, are taken up as well as pointers to the great philosophers and philosophical schools. Try to synthesize just some of this, say focusing on the original discussion paper and a few selected commentaries.
I assume that you know the truly basic set theoretical constructions (say from “MatM”). Therefore, we limit the discussion here to a set of exercises meant to test this. You will also meet some extensions of known concepts and results. Also, you find elements of the exercises where you have to rely on more than just naive set theory (e.g. on familiarity with cardinal numbers).

Exercises

2.1 List the major binary (or unary) set theoretical operations. Give an illustration of these operations.

2.2 Use basic set-theoretical constructs, especially that of an ordered pair, to define the important notion of a function, often also called a map or a mapping. Also define the following notions, introducing extra objects where needed for the notions to make sense: injective function, everywhere defined function, surjective function, bijective function, bijection, embedding (da.: “indelejring”).

2.3 Explain what is meant by a right inverse and by a left inverse of a given function, \( f \). What is required of \( f \) for these objects to exist? Try and illustrate these notions, either by some kind of figure or by (one or more) diagrams.

2.4 The standard version of the distributive law is \( A \cap (B \cup C) = (A \cap B) \cup (A \cap C) \). This is the simplest version of a very useful fact of set theory which transforms an intersection of unions to a union of intersections. A much more general version of this fact is the following result: Let \( I \) be a set (an index set) and let \( J = (J_i)_{i \in I} \) denote a family of sets (index sets) indexed by
i ∈ I. For each i ∈ I, let \((A_{i,j})_{j \in J_i}\) be a family of sets indexed by \(j \in J_i\). Denote by \(J^I\) the set of all functions \(f\) defined on \(I\) for which \(f(i) \in J_i\) for each \(i \in I\). Then

\[
(*) \quad \bigcap_{i \in I} \bigcup_{j \in J_i} A_{i,j} = \bigcup_{f \in J^I} \bigcap_{i \in I} A_{i,f(i)}.
\]

(i) By choosing \(I\) and the family \(J\) of index sets appropriately, show how the standard distributive law can be obtained as a very special case of (*).

(ii) Prove (*). Note: The reader may find it a lot easier first to do (iii) below.

(iii) The requested proof under (ii) may be facilitated by a drawing which helps to keep track of the structure. This is most simply done in case all the index sets \(J_i\) is one and the same set, say \(J\). Then the notation chosen in (*) agrees with standard notation in the sense that the standard notation for the set of functions defined on \(I\) and with values in \(J\) is \(J^I\). The visualization we have in mind which is associated with (*) then views the given data as the matrix with rows indexed by \(i \in I\) and columns indexed by \(j \in J\) for which the \((i,j)\)’th entry is the set \(A_{i,j}\). Make this visualization clear and indicate how one can visualize both sides of (*).

(iv) Using de-Morgans rules, formulate and prove a “dual” version of (*) which transforms a union of intersections into a certain intersection of unions.

(v) A paving \(\mathcal{E}\) in a set \(X\) is a non-empty collection of subsets of \(X\), i.e. \(\mathcal{E} \subseteq \mathcal{P}(X)\) and \(\mathcal{E} \neq \emptyset\) (here \(\mathcal{P}(X)\) denotes, as usual, the power set of \(X\), i.e. the paving of all subsets of \(X\)). A topology in the set \(X\) is a paving containing \(\emptyset\) and \(X\) which is closed under finite intersections (if \(E_i \in \mathcal{E}\) for all \(i \in I\) and if \(I\) is finite then \(\bigcap_{i \in I} E_i \in \mathcal{E}\)) and arbitrary unions (if \(E_i \in \mathcal{E}\) for all \(i \in I\) and \(I\) is any set, then \(\bigcup_{i \in I} E_i \in \mathcal{E}\)). Let \(\mathcal{E}\) be a paving on \(X\) and assume that \(\mathcal{E}\) contains the empty set as well as the set \(X\). Denote by \(\mathcal{E}^\tau\) the paving of arbitrary unions of finite intersections of sets in \(\mathcal{E}\). Show that this paving can be characterized as the smallest topology containing \(\mathcal{E}\). This paving is called the topology generated by \(\mathcal{E}\).

(vi) A lattice in a set \(X\) is a paving which is closed under finite intersections and finite unions. In the spirit of (v) show that the lattice generated by a paving \(\mathcal{E}\) can be obtained either as
Basic constructs of naive set theory

the paving of finite intersections of finite unions of sets in \( E \) or, alternatively, as the paving of finite unions of finite intersections of sets in \( E \). Note: Again, this is quite easy using the distributive law(s). However, simple closure results as this one and as the result under (v) are not always obtainable. For instance, if we ask for the \( \sigma \)-lattice generated by a paving, where the “\( \sigma \)” refers to a requirement that now we consider closure properties under countable set-operations rather than under finite operations, then the situation is more complicated (involves ordinal numbers up to \( \omega_1 \) as we may return to in another exercise). As a result, most people find it more difficult to come to grips with the basic sets of measure theory whereas basic sets in general topology can normally be written down explicitly.

Note 1: A convenient notation is introduced for basic set-theoretical operations applied to a paving \( E \) in a set \( X \). Basically, a paving is \( \tau \)-closed if it is closed under the operations listed in \( \tau \) and, for a paving \( E \), the \( \tau \)-closure of \( E \), denoted \( \text{cl}(\tau)E \), is the smallest paving containing \( E \) which is closed under the operations signalled in \( \tau \). Typically, \( \tau \) consists of a list combining signs for set-theoretical operations (such as \( \cup \), \( \cap \), \( \Delta \), \( \setminus \), \( \overline{\cdot} \): union, intersection, symmetric difference, difference, complement) with cardinality restrictions regarding \( \cup \) and \( \cap \), either “f” for finite, “c” for countable and “a” for arbitrary. In this notation, a topology in \( X \) is a \((\emptyset, X, \cup, a, \cap f)\)-closed paving in \( X \) and the topology generated by \( E \) is the paving \( \text{cl}(\emptyset, X, \cup, a, \cap f)E \); and a Borel structure, also called a \( \sigma \)-algebra, is a \((\overline{\cdot}, \cup c)\)-closed paving and the Borel structure or \( \sigma \)-algebra generated by \( E \) is the paving \( \text{cl}(\overline{\cdot}, \cup c)E \). Unions and intersections are often placed as subscripts, e.g. \( E_{\cap f} \) denotes the paving of finite intersections of sets in \( E \). Clearly, \( E_{\cap f} = \text{cl}(\cap f)E \). Notation such as \( E_{\cap f, \cup c} \) means the paving of countable unions of finite intersections of sets in \( E \), “incidentally” equal to \( \text{cl}(\cap f, \cup c)E \). But warning, note that for instance \( E_{\cap f, \cup c} \) is usually far from \( \text{cl}(\cap f, \cup c)E \). It is costumary to write \( \text{co}(E) \) rather than \( E_{\overline{\cdot}} \), the paving of complements of sets in \( E \).

Usually, \( \text{cl}(\tau)E \) exists for every paving. The trivial proof of this goes as follows: “Clearly”, \( \mathcal{P}(X) \) is \( \tau \)-closed and contains \( E \) and, just as “clearly”, the intersection of all \( \tau \)-closed pavings containing \( E \) is itself \( \tau \)-closed. Voila! It is important to realize that this proof is unsatisfactory in spite of its simplicity. The
proof says nothing about which sets to include in the $\tau$-closure. The exercise points to some instances where we can do better and identify exactly which sets to include in the $\tau$-closure. In fact, the content of (v) and (vi) can be formulated as follows:

$$cl(\emptyset, X, \cup f) E = E \cap f, \cup a$$

$$cl(\cup f, \cap f) E = E \cup f, \cap f = E \cap f, \cup f.$$  

Note 2: As most (all?) readers of these notes will know, the kind of constructions discussed in this exercise are important for the beginnings of general topology and measure theory. The constructions are also important for a study of the nature of subsets of standard sets such as the reals. But for such a study further set theoretical operations become important, especially two types of operations, viz. projections (e.g., which sets do you obtain by projecting “nice” sets in the plane, say Borel sets, onto a line ? (in fact you may obtain other sets than Borel subsets of the line!)) and then there is one very special operation, called the Souslin operation, which is related to both union and intersection and is a kind of hybrid operation being in one sense countable and in another more than countable. Given more time for the course, it would be a natural ground to enter. The next exercise scratches the surface by giving the definition of the Souslin operation and some indication of what can be obtained using it.

2.5 Let $N^N$ be the set of sequences $n = (n_1, n_2, \cdots)$ of natural numbers, let, for $k \geq 0$, $N^k$ be the corresponding set of sequences of length $k$ (with the understanding that $N^0 = \{\emptyset\}$) and denote by $n \preceq n|k$ the projection of $N^N$ on $N^k$. A tree is an ordered set such that every left section is well-ordered (the well orderings appearing below are just finite sections of $N$). A rooted tree is a tree with a smallest element. The Souslin tree is the rooted tree denoted $N(N)$ and defined by

$$N(N) = \bigcup_{k=0}^{\infty} N^k$$

in the natural ordering, i.e. $n \preceq m$ means that $n$ is a restriction of $m$ (including the case $n = m$).

Let $X$ be a set and $E$ a paving in $X$. A Souslin scheme $A$ over $E$ or an $E$-Souslin scheme is a map $A : n \preceq A(n)$ of $N(N)$.
into $\mathcal{E}$, except that we insist that $A(\emptyset) = X$ (even if $X \notin \mathcal{E}$). The associated Souslin set, $S(A)$ is the set

$$S(A) = \bigcup_{n \in \mathbb{N}} \bigcap_{k \geq 0} A(n|k).$$

The set-theoretical operation $A \curvearrowright S(A)$ is called the Souslin operation. By $S(\mathcal{E})$ we denote the paving of all sets $S(A)$ with $A$ an $\mathcal{E}$-Souslin scheme.

(i) Prove that $S(\mathcal{E})$ is $(\cup^c, \cap^c)$-closed. **Hint:** Really, closure under $\cup^c$ is quite easy to establish, though it does need a little creativity on your side. As for closure under $\cap^c$, consider $\mathcal{E}$-Souslin schemes $A_1, A_2, \cdots$. You have to define an $\mathcal{E}$-Souslin scheme $A$ such that $S(A) = S(A_1) \cap S(A_2) \cap \cdots$. The idea is to mix the given Souslin schemes so that at level $k$ (i.e. for $n \in \mathbb{N}^k$) you only find sets of the form $A_\nu(m)$ with $m$'s in $\mathbb{N}^l$ for $\nu$ and $l$ fixed (but depending on $k$). The association $k \curvearrowright (\nu, l)$ could be $1, 2, 3, 4, 5, 6, \cdots \curvearrowright (1, 1), (2, 1), (1, 2), (3, 1), (2, 2), (1, 3), \cdots$. Then define the desired Souslin scheme by induction as briefly indicated here: To define the $A(n) = A(n_1, \cdots, n_k)$'s first determine the associated $(\nu, l)$. If $l = 1$, place $A_\nu(1), A_\nu(2), \cdots$ at the new nodes (with repetition for each $(n_1, \cdots, n_{k-1})$), and if $l > 1$, do something a bit similar, paying attention to the assignements made last time you chose sets from the Souslin scheme $A_\nu$.

**Note:** By more elaborate constructions (or otherwise) one can prove that $S(\mathcal{E})$ is even closed under the Souslin operation.

(ii) Assume that every complement of a set in $\mathcal{E}$ is a countable union of sets in $\mathcal{E}$ or, more generally, that $co(\mathcal{E}) \subseteq S(\mathcal{E})$. Prove that then $S(\mathcal{E})$ contains $cl_{\mathcal{E}, \cup^c} \mathcal{E}$, the $\sigma$-algebra generated by $\mathcal{E}$. **Hint:** Look at $S(\mathcal{E}) \cap co(S(\mathcal{E}))$.

(iii) For the case $X = \mathbb{R}$ and $\mathcal{E}$ the paving of closed intervals in $\mathbb{R}$ show that every Borel set (recall, a set in the $\sigma$-algebra generated by the paving of open sets) can be represented as an $\mathcal{E}$-Souslin set. Illustrate this by representing the Cantor set in this way. Use the fact to determine the cardinality of the set of Borel sets in $\mathbb{R}$. Generalize if you feel like it (to $\mathbb{R}^n$, to separable metrizable spaces, ...).

$\dagger$ interested readers can request a copy of a manuscript on these issues from me
This chapter is a key to much of what is to follow, especially to axiomatic set theory. The infinite has been an object of fascination through times, surely dating back beyond history to prehistoric man. With its dimension of the unreachable and its exposure of man's limitations to the finite, speculations over the infinite invites to religious irrational thinking. Perhaps it was for this reason, the old Greeks thought it better to exclude the infinite (the actual infinite) from their philosophical universe, yet admitting the striving towards the infinite (the potentially infinite). Exceptions to this attitude existed, notably presented by one of the greatest scientists of all times, Archimedes (287-212).

In modern mathematics we must insist that the notion of infinity has a firm place which enables sound reasoning. How else can we deal in precise terms with disciplines such as the infinitesimal calculus emerging through works of Newton (1642-1727), Leibniz (1646-1716) and their followers?

How do we go about the problem, allowing what we cannot reach to enter into our mathematical universe? We shall start with the beginning, the early encounters with the infinite, then quickly turn to illustrations which serve to adjust our preconceptions. Further material in the chapter may well appear quite technical to the reader. Still, we stay entirely at the naive level. Not until Chapter xx, when axiomatic set theory is at our disposal, can we introduce infinity in a rigorous manner. But the way will be paved with this chapter!
3.1 The child

All of us have met the infinite in our childhood. Many for the first time watching the sky on a clear night, wondering about what is out there. How many stars? And just how much of “it” is there? Does it extend endlessly or is it limited? And will it continue to be there and has it always been so? Fascinating!

You may also have competed with friends on just how many times you can fold a piece of paper or halve a branch, cut a layer cake or what the case may be, and speculated if, in principle this could be done infinitely many times. Or one day at the beach we wonder just how many grains of sand there are†. Surely, to us uncountably many but, in reality, only finitely many one should think. Or would our reflections tempt us to use the term “infinite” in situations where we cannot count. This is certainly done in a great number of cases, e.g. in modelling based on the differential calculus. In such cases one is well aware of the essential finite nature of the object of the modelling – think of statistical thermodynamics, for example – and only sees the analysis based on infinite processes as a pragmatic way to usable results which cannot be obtained otherwise. Let us agree that the essence of the infinite is not just “many” but the potential impossibility of counting.

To sum up: Consciously or unconsciously we have all been confronted with the infinite in various disguises in our youth or even in early childhood.

3.2 Comprehension

Now that we have realized that infinite objects have a place in our mind, we want to investigate closer the peculiarities of the infinite. It is only natural to look at the simplest infinite object we can think of. The set \( \mathbb{N} \) of natural numbers comes to mind:

\[
\mathbb{N} = \{1, 2, 3, \cdots\}.
\]

The problem is of course, exactly what does the “dot-dot-dot” stand for? Let us start with what we do know: the beginning 1, 2 and 3. These numbers – and we could have continued the list a bit – have a firm place already in our mathematical universe and have since long been accepted

† Do you belong to those who have wondered about the myriads of grains of sand? You are not the first. Archimedes proceeded you. He set out to define a concrete natural number which is larger than the number of sand grains, had the entire universe been filled with sand!
as objects with meaning and objects we can manipulate and work with in many connections. Let us, for the time being stick to this view. Later, cf. Sectionxx.x, we shall also see how one can introduce numbers such as 1, 2 and 3, should we doubt what they “really” are.

So, for now we accept that each individual natural number has a firm place in our mathematical universe. From there to the set of natural numbers – an infinite set! – there is in a sense not far. We just think about it! To our mind it is the totality of all natural numbers. It may be helpful to imagine that you list on the blackboard or elsewhere each one of the natural numbers and then step back a bit and look at what you have: The set of all natural numbers. This is what we did in (3.1). The two braces represent the process of comprehension – putting into our mind – and what is between the braces is what is being comprehended.

| The idea of comprehension | Everything we can think about in precise terms has a place in the mathematical universe. |

Expressed briefly, the idea is the optimistic “I think about it, therefore it is”.

According to this principle the natural numbers as such, the set of natural numbers, is a well defined mathematical object. However, the principle says nothing about how exactly this set is defined. And for good reasons. Accepting the idea of comprehension as a generally valid principle leads to contradictions as we shall see when discussing Russell’s paradox in Sectionxx.x. But as a principle worth striving for, the idea of comprehension is important and a good guide later on when we create an axiomatic basis for set theory. For now we just accept the idea when applied to the totality of natural numbers and take it for granted that the set of natural numbers is a well defined object. The way we think about this object is – of course – by comprehension. After all, there was no difficulty putting it into your mind, was there?

3.3 The hotel

The strange world of the infinite was introduced into rigorous mathematics by Cantor (1845-1918). With the work of Cantor and his followers a sound basis for all – well, almost all – of mathematics emerged. Not everyone understood at the time the scope of the new creation but gradually the ideas caught on and they are now universally accepted by mathematicians. One of the most enthusiastic supporters and promoters
was Hilbert (1862-1943) who stated that “no one shall expel us from the paradise that Cantor has created”. In his lectures from around 1920, Hilbert presented a mathematical thought experiment, *Hilbert’s Hotel*, intended to illustrate some of the characteristics of infinite sets.

In Figure xx we see a principal sketch of the hotel with the reception and then a long corridor leading to the rooms, each identified (“coded”) by a natural number. And all natural numbers are used as the hotel stretches into the infinite with infinitely many rooms.

**Stories about Hilbert’s Hotel** We shall tell only three stories.

1. **story (adding one)** On a dark and stormy night an exhausted traveler arrives at Hilbert’s Hotel and seeks shelter for the night. Unfortunately, all rooms are occupied. However, the receptionist calls the guests over the microphone system: “The receptionist to all our guests: Please leave your room and move to the adjacent room further down the corridor”. With this manoeuvre, room 1 becomes free and the traveler can be accommodated there.

Thus, we have an instance where one of Euclid’s (around 300 B.C.) axioms (taken from his group of “common notions”), viz. “whole is greater than the part” does not hold. There is nothing contradictory as such about it. It is just one of the strange features we have to accept when we allow infinite objects in our universe. We may say that adding an element (the traveler) to the original set (consisting of the original guests) does not change anything in a sense, since after a simple regrouping† the new set with the traveler added looks exactly as the original set. Or we may say that removing an element from the totality of all guests does not really change anything.

2. **nd story (adding a replica)** You should know that the popular hotel was soon copied by enterprising businessmen who set up a true copy of Hilbert’s hotel, Klamphuggernes Hotel. This too became popular and was always fully occupied as was the original Hilbert’s Hotel. Then, the second story goes: On a dark and stormy night, what was put up in haste fell asunder. So all guests at Klamphuggernes Hotel were in need for shelter for the night. They put their hopes to the receptionist at Hilbert’s Hotel. And, indeed, he felt pity for the many people in distress and managed to create enough vacant rooms by calling: “The receptionist to all our guests: Please move to the room with room number the double of yours”. This manoeuvre solved the problem as you will easily realize.

† well, admittedly not that simple if you insist on speculating on just how, physically, the regrouping is effectuated and do not accept the hotel as the thought experiment it is intended to be
So even adding an infinite set to an infinite set or taking away an infinite set from an infinite set need not really change anything.

**3.3 The hotel**

**3.rd story (the galaxy version)** Before telling the story, you should know that the fully expanded universe consists of infinitely many galaxies, each containing infinitely many stars, each having infinitely many planets in orbit around it and each planet with infinitely many cities, each having infinitely many streets, each with infinitely many street numbers, each number with infinitely many storeys and each storey with infinitely many rooms. Each room is occupied by a human being. On one of the planets (Earth) associated to one of the stars (the Sun) in one of the galaxies (the Milky Way) one finds good old Hilbert’s Hotel.

Now then, our final story: On a dark and stormy night – well, **ragnarok**† in fact – the universe collapsed, except, of course, solid built Hilbert’s hotel. All were in distress. The universal idea occurred to all, that the only possibility for rescue from destruction was to seek shelter in Hilbert’s Hotel. So the inhabitants of the universe went there, some traveling short distance only, but most traveling in ultra fast space crafts.

The receptionist at Hilbert’s Hotel had noted that this night was really worse than anything encountered before and understood instantly what his task was.

... Here the basic story ends – but not the mathematics connected with it. The task for the receptionist is not that easy this time, but actually he can accommodate everybody provided the use of the phrase “infinitely many” which occurred eight times in the story means “which can be numbered just as is the case with the rooms in Hilbert’s Hotel”.

This means that each inhabitant in the fully expanded universe can be identified by an eight-tuple:

\[(n_1, n_2, n_3, n_4, n_5, n_6, n_7, n_8)\]

of natural numbers. The problem then is really this: Can we assign a room (i.e. in effect a natural number) to every eight-tuple such that no two eight-tuples are assigned the same room.

Often, when confronted with a problem of a nature one has not met before, it is helpful to change the actual problem at hand and consider the simplest instance of a problem which has enough of the flavour of the original problem in it to reflect back and illuminate the original problem.

† the final day in Nordic mythology
The infinite

– and yet is simple enough to enable you to say something about it, at best to solve the simplified problem completely. This strategy works well here. As simplified problem we consider a situation with only two levels instead of eight. The simplified problem then is if we can accommodate every pair \((n_1, n_2)\) of natural numbers in Hilbert’s hotel. And indeed we can. Figure xx indicates how. Start with \((1, 1)\) and allocate room number 1 to that pair, then proceed along the arrows when allocating subsequent rooms. Eventually (?), all rooms are allocated to pairs of natural numbers and every such pair has been accommodated in a single room.

The exploitation of the result about pairs just established to solve the original problem regarding eight-tuples is left to the reader as an easy exercise Exercise ??

3.4 The ballroom

It looks as if any infinite set can be accommodated in Hilbert’s Hotel. But is this really so? We can think of two possible scenarios. One possibility is that the infinite is an absolute. There is only one type of infinite and this can be represented by the set of natural numbers or the rooms in Hilbert’s Hotel, if you wish. This has the attraction to point to a theory of great unification. Only one kind of infinity which can, of course be represented in various disguises. But, for another scenario, we could argue that for mathematics to be flexible it is more attractive with several types of infinity at our disposal. That will give greater freedom and result in greater expressive power. It turns out that the latter scenario is the one to opt for. If we allow even one kind of infinity into our mathematical universe, we better allow several.

The possibility of several infinities makes Hilbert’s Hotel less suitable as a playground for our further discussions of the infinite. Instead, we turn to Jessen’s Ballroom as the appealing device intended to support intuition. Jessen (1907-1993), prominent Danish mathematician, was renowned for his elegant and meticulously prepared lectures. He was a great admirer of Hilbert and in his lectures about Cantors set theory, he too suggested a mathematical thought experiment of great suggestive power. The thought experiment we call Jessen’s Ballroom. It is intended as a visual aid and guide for us when attempting to compare infinities. It consists of two long (very long!) rows, one for the boys, the other for the girls. The rows flank the vast dance floor. At the entrance we find the master of the ballroom – be it Cantor or Jessen – who always
ensures that the proper order is kept: “Boys to the right, girls to the left, boys to the right, girls to the left ...”.

Thus two sets are involved, $X$ as the set of boys and $Y$ as the set of girls. A typical element of $X$ – a boy then – is denoted $x$, whereas a typical element of $Y$ is denoted $y$. Our task is to investigate if there are more girls than boys or the other way round – or perhaps there are just as many girls as boys. We are most interested in situations with $X$ and $Y$ infinite, but what we will say below applies quite generally to any kind of sets, finite or not.

Firstly we shall agree on precise definitions enabling the kind of comparison of sets we are looking for. As a starter, consider the following problem: “Which of your hands hold most fingers?” A stupid question, you may react. Of course there are equally many fingers on the two hands. And if you are pressed to explain in greater detail why this is so, one possibility is that you start counting: “One, two, three, four, five on that hand and one, two, three, four, five on the other, equally many, so there you are”. But this process of counting is perhaps too complicated and not at all well defined when it comes to more complicated sets, to infinite sets. When we discussed various sets in connection with Hilbert’s Hotel that was not what we did. Thinking about it, you realize that you can do better than counting. It is simpler to pair the fingers on the two hands as in Figure xx and this process of pairing makes sense, also for “large” sets.

Hence, let us agree as follows:
Definition 3.4.1. Two sets $X$ and $Y$ are considered of equal size in a purely set theoretical sense, and we say that the sets are *equipotent* and write $|X| = |Y|$ if there is a complete pairing of the boys with the girls so that when the pairs are formed and the dance begins, *all* are dancing.

One way to reformulate this emerges if we view the situation from the point of view of the boys and assume that the pairs are formed by the boys inviting the girls for dance. The condition of the complete pairing then means that this can be done in such a way that the ensuing map $x \mapsto y$ of $X$ into $Y$ (with $y$ the girl invited for dance by the boy $x$) is a bijection. Therefore,

**Theorem 3.4.2.** Two sets $X$ and $Y$ are equipotent if and only if there exists a bijection of $X$ onto $Y$.

When there exists no bijection between $X$ and $Y$, the sets are not equipotent and we write $|X| \neq |Y|$.

**Example 3.4.3.** Assume that $X = Y = \mathbb{N}$. Or you may say that $X$ and $Y$ are different sets with the set of boys, $X$ being such that the elements can be numbered by a bijection between $X$ and $\mathbb{N}$ and, similarly, the elements of $Y$ are numbered by a bijection between $Y$ and $\mathbb{N}$. Below we largely ignore the bijections referred to and identify $X$ as well as $Y$ with $\mathbb{N}$.

Now assume that any boy $x$ invites the girl $y = x + 1$ for dance. This will result in all boys dancing, but there will be one wallflower among the girls, viz the girl 1. No one has invited her for dance.

But by inviting the girls for dance in a more accommodating way, say such that the boy $x$ invites the girl $y = x$ for dance, a complete pairing results and hence the two sets are equipotent. Of course – after all, they are just two copies of one and the same set.

We then fix what it shall mean that there are at least as many girls as boys. Again, we agree to view it from the point of view of the boys:

**Definition 3.4.4.** We say that there are at least as many girls as boys, and write $|X| \leq |Y|$, if the boys can invite the girls to dance in such a way that all boys dance - but there may be wallflowers among the girls.

Just as with Theorem 3.4.2 we realize that the definition can be expressed using well known terminology for mappings:

**Theorem 3.4.5.** We have $|X| \leq |Y|$ if and only if there exists an injective map of $X$ into $Y$. 

The ballroom

3.4 The ballroom

In other words, cf. Section xx.x, the condition is that \( X \) can be embedded into \( Y \). Possible notation: There exists a map \( X \hookrightarrow Y \).

Our final definition is really the one we are after when we ask if Hilbert’s Hotel is large enough to accommodate all kinds of infinities. The definition relies on the two previous ones and reads as follows:

**Definition 3.4.6.** We say that there are effectively more girls than boys, and we write \( |X| < |Y| \) if \( |X| \leq |Y| \) holds and \( |X| = |Y| \) does not hold.

In other words, there are effectively more girls than boys if the boys can invite the girls for dance in such a way that all boys dance but, no matter how they invite the girls for dance there will always be wallflowers among the girls.

The importance of the above definitions is that they express our intentions and natural instincts regarding the meaning of “equally many”, “at least as many” or “effectively more”.

The discussion should have convinced you of the good sense of the above definitions. Have we cheated you by hiding not-so-obvious elements? Not really, but we already now point to one essential feature of the definitions: They all depend on the notion of mapping (or function). And this turns out to be a not-so-obvious notion. Fact is that an expression as, say “consider the set of all functions of \( X \) into \( Y \)” is a very complicated statement. Though we can define what a function is by describing what we mean – and this is the level we will be satisfied with through almost all of this book – we cannot get our fingers at this notion in more constructive terms. Thus the definitions given depend on the notion of a function. Later, cf. Section xx.x this will give room to the odd situation that the notion of function can be changed and lead to different outcomes for the definitions above.

Another comment concerns the use of the notation \(|X|, |Y|\) etc. For the time being the notation is only used as a device to abbreviate statements. Thus \(|X|\) will not appear by itself, but only in connection with other instances of this notation in accordance with the definitions above. This situation will be changed in Section xx.x, where we shall attach a precise meaning for any set \( X \) to \(|X|\).

Now then, let us search for a set which is effectively larger than \( N \), i.e. a set that cannot be accommodated in Hilbert’s Hotel. The set of \( N \)-tuples of natural numbers was too small however large \( N \in \mathbb{N} \) is (see Exercise (viii)). What about the set of infinity-tuples? In fact, this will work, and this is so even if we only allow a choice among two
numbers at each position. Let us choose these two numbers as 0 and 1.
To be precise, the set we are considering is the set $2^\mathbb{N}$ of all sequences
(infinity-tuples) $(\varepsilon_1, \varepsilon_2, \varepsilon_3 \cdots)$ of 0’s and 1’s.

**Theorem 3.4.7** (Cantor’s theorem). $|2^\mathbb{N}| > |\mathbb{N}|$.

**Proof.** Clearly, $|2^\mathbb{N}| \geq |\mathbb{N}|$. Taking $\mathbb{N}$ as the set of boys and $2^\mathbb{N}$ as the
set of girls, we shall now show that however the boys invite the girls for
dance, there is a wallflower among the girls. Let

\[
(\varepsilon_{1,1}, \varepsilon_{1,2}, \varepsilon_{1,3}, \cdots) \\
(\varepsilon_{2,1}, \varepsilon_{2,2}, \varepsilon_{2,3}, \cdots) \\
(\varepsilon_{3,1}, \varepsilon_{3,2}, \varepsilon_{3,3}, \cdots) \\
\vdots \\
\vdots
\]

be the list of partners chosen by the boys. Clearly then,

\[
(\varepsilon_1, \varepsilon_2, \varepsilon_3, \cdots)
\]

with the $\varepsilon$’s defined by

\[
\varepsilon_n = 1 - \varepsilon_{n,n} \text{ for } n \in \mathbb{N}
\]

is a wallflower as she is not dancing with the first boy (is she?), not with
the second either, or the third etc. \(\square\)

The elegant proof was discovered by Cantor around 189x and is known
as *Cantor’s diagonal argument* (by modifying the diagonal elements in
the list occurring in the proof you get at a wallflower).

So there are at least two types of infinities. In fact there are infinitely
many types of infinities since, for any set, we can construct another
which is effectively larger:

**Theorem 3.4.8** (Cantor’s theorem, general case). For any set $X$, $|\mathcal{P}(X)| > |X|$.

**Proof.** Clearly, $|\mathcal{P}(X)| \geq |X|$.

Now take $X$ as the set of boys and $\mathcal{P}(X)$ as the set of girls and consider
the dance one evening in Jessen’s Ballroom with all boys on the floor.
For any boy $x$ we denote his partner by $A_x$. Then the girl

\[
B = \{x | x \notin A_x\}
\]
is a wallflower! Indeed, no boy $x_0$ is weird enough to have chosen this girl as his partner since then $A_{x_0} = B$ would hold, hence, for every $x \in X$ the bi-implication $x \in A_{x_0} \iff x \in B$ holds true and, taking $x = x_0$, this leads to the absurd statement $x_0 \in A_{x_0} \iff x_0 \not\in A_{x_0}$. □

It is an important fact that one may represent one of the infinities effectively larger than the infinity of $\mathbb{N}$ by a familiar set:

**Theorem 3.4.9.** The set of reals is effectively larger than the set of natural numbers: $|\mathbb{R}| > |\mathbb{N}|$.

For a proof, see Exercise (xii).

### 3.5 Properties of equipotence

We start with a very useful and important result:

**Theorem 3.5.1** (Schröder Bernstein’s theorem). If, for two sets $X$ and $Y$, both $|X| \leq |Y|$ and $|Y| \leq |X|$ hold, then $|X| = |Y|$.

This seemingly obvious result is not trivial. We shall give a proof formulated in the terminology of Jessen’s Ballroom.

**Proof.** We have to construct a pairing of the boys (elements of $X$) with the girls (elements of $Y$). What we know is that the boys can invite the girls for dance such that all boys dance and that the girls can invite the boys for dance such that all girls dance. Let us have two fixed choices of partners in mind, one for which all boys dance, and one for which all girls dance. Assume that a dance is organized every evening, the first with the boys inviting the girls for dance, the second with the girls inviting the boys for dance and so on alternating with the boys and the girls choosing their partners. We emphasize that the boys always use the same strategy when they ask the girls for dance. So each boy has one favourite partner among the girls and this partner never changes. Likewise, each girl has a fixed favourite boy whom they always ask for dance, giving the chance, i.e. when it is the girls turn to ask for dance.

We also imagine that communication among the girls and boys is very lively, but only while they dance and only among partners. We imagine that the dances have been going on “forever” (through evenings number 1, 2, 3, etc.).

The complete pairing of boys and girls is now effectuated in the following way: First, all girls who are either wallflowers when the boys invite for dance or else know about such wallflowers are asked to invite
their preferred partner for dance. Secondly, all remaining boys (if any) are asked to invite their favourite partner for dance. We claim that in this way a complete pairing results so that all – boys as well as girls – are dancing.

To clarify the first step in the construction, let us imagine that one evening Charles invites Beth for dance and during the dance Charles tells Beth that the evening before he was dancing with Dorothy who told him that no one had asked her for dance the evening before that. So now Alice knows that there is a wallflower among the girls, Dorothy. So in our construction, Alice will, in the first step, invite her preferred partner, Andrew, for dance.

In the second step, any remaining boy, say Jim, will find that his preferred partner, say Jane, is not already dancing. In fact, if this was not the case, we note that since Jane is not herself a wallflower, she could only have heard about wallflowers among the girls from Jim and he in turn must have heard it from his admirer, say Karen. So Karen would know about wallflowers among the girls and would, therefore, have asked Jim for dance in the first step. This contradicts the fact that Jim is one of the remaining boys. We must conclude that his preferred partner, Jane is still available and so he hastens to ask her for dance in the second step.

So, after the second step, all boys are dancing. But also all girls are dancing: Those who know about wallflowers among the girls already entered the dance floor in the first step and if a girl, say Sue does not know about wallflowers among the girls and must then have an admirer, say Sam and then a moments consideration tell you that Sam invited her for dance in the second step. So the pairing is complete.

We now use Schroder Bernstein’s theorem to identify a very well known standard set which is equipotent with $\mathcal{P}(\mathbb{N})$:

**Theorem 3.5.2.** The set $\mathbb{R}$ of real numbers is equipotent with $\mathcal{P}(\mathbb{N})$, the power set of $\mathbb{N}$.

**Proof.** Note that $t \hookrightarrow \{ q \in \mathbb{Q} | q \leq t \}$ is an embedding of $\mathbb{R}$ into $\mathcal{P}(\mathbb{Q})$ and that $A \hookrightarrow \sum_{k \in A} 10^{-k}$ is an embedding of $\mathcal{P}(\mathbb{N})$ into $\mathbb{R}$ and apply Schr oder Bernstein’s theorem.

In the above proof we used without saying so explicitly the fact that when two sets – such as $\mathbb{Q}$ and $\mathbb{N}$ above – are equipotent, then so are their power sets. This and similar obvious facts we collect in one theorem.
3.6 Counting to infinity

Theorem 3.5.3. Assume that \((X, X'), (Y, Y')\) and \((Z, Z')\) are pairs of equipotent sets, and, for the case of (iii) below that \(X \cap Y = X' \cap Y' = \emptyset\). Then the following properties hold:

\[
\begin{align*}
|\mathcal{P}(X)| & \equiv |\mathcal{P}(X')| \quad (3.2) \\
|X \cup Y| & \equiv |X' \cup Y'| \quad (3.3) \\
|X \times Y| & \equiv |X' \times Y'| \quad (3.4) \\
|X^Y| & \equiv |(X')^{Y'}|. \quad (3.5)
\end{align*}
\]

The simple proof is relegated to the exercises.

3.6 Counting to infinity

We know how to compare two infinite sets. Now we shall investigate, given just one infinite set, if we can somehow count its elements. In practical terms this is impossible but in our abstract mathematical universe this is conceivable. The idea is in a sense the same as for finite sets. If, for example \(X\) is a set with 5 elements, we can count its elements by singling out a 1.st element, a 2.nd element, a 3.rd element, a 4.th element and a 5.th element. This counting gives a succession or an order of the elements. When we think about 1, 2, 3, 4, 5 (and larger numbers) in this way, we talk about ordinal numbers. The idea then is to continue until we have exhausted the elements of any given set \(X\), even if \(X\) is infinite. In this way we will be led to ordinal numbers going beyond the usual finite ordinals. The new ordinal numbers are the transfinite ordinals.

So, after all we are giving in to the primitive counting of fingers as first suggested when asking which hand holds most fingers. Instead of the elegant method of pairing, we return to the much more cumbersome method of counting. This is needed in order to better understand the nature of the infinite. Using only pairing, we could say that one set is larger, in the sense of equipotence, than the other. But a more detailed comparison requires that we understand more about the structure of infinite sets. But there is no structure in a set, you may argue. A set is just an abstract collection of elements. Right you are. What we shall do is to introduce structure in a set under study. The notion of well-ordering introduced in Section xx turns out to be just the right tool. We shall rely on Zermelo’s well ordering theorem, already formulated in Section xx.

In order to build the readers intuition about the ordinal numbers we
consider a “huge” infinite set $X$ and a well-ordering on $X$. We assume that the well-ordering has a largest element (cf. Exercise xx).

As $X \neq \emptyset$ ($X$ is huge, definitely not empty!), there exists a first element in $X$. Let us denote that element by $0$. Remove it from $X$. What remains is $X \setminus \{0\}$, a non-empty set. Call the first element in that set $1$. What we did was to look at the successor of $0$: $1 = 0^+$. Continue in this way and define $2 = 1^+$, $3 = 2^+$ etc. In this way we will have constructed elements

$$
\begin{array}{cccccc}
0 & 1 & 2 & 3 & 4 & 5 \\
\end{array}
$$

To be sure, we ought to have written $x_0, x_1, x_2, \cdots$ in place of $0, 1, 2, \cdots$. However, we find it convenient and suggestive of what we are aiming at to use the familiar notation $0, 1, 2, \cdots$.

Now remove all elements $0, 1, 2, \cdots$ from $X$. As $X$ is large, in particular uncountable, the set that remains is still non-empty. As such, it has a first element, which we denote $\omega$. Then we are this far in our attempt to count the elements of $X$:

$$
\begin{array}{cccccc}
0 & 1 & 2 & 3 & 4 & 5 & \omega \\
\end{array}
$$

The element $\omega$ is interesting. It is the first element with more than finitely many preceding elements. The left section $V(\omega)$ can be used as a model of the natural numbers $\mathbb{N}_0$. We can also characterize $\omega$ as the first limit element, the first one after 0 which is not a successor, cf. section xx. After $\omega$ comes elements $\omega + 1$, $\omega + 2$ etc. When we are finished with these elements, we have constructed one more “$\omega$-block”:

$$
\begin{array}{ccccccc}
0 & 1 & 2 & 3 & \omega & \omega + 1 & \omega + 2 & \cdots \\
\hline
\omega\text{-block} & \omega\text{-block} \\
\end{array}
$$

Now comes the element we denote $2\omega$ (later on it would be more correct to call this element $\omega 2$). This is the second limit element. We continue by listing elements of $X$, removing them, listing new elements etc. So we encounter a new $\omega$-block starting with $2\omega$, then another starting with $3\omega$ and so on. Then we are this far:
But it is far until $X$ is exhausted. Still, we have only removed a countable set from $X$. Many elements remain. The first one is denoted $\omega^2$. We are now this far:

We realize that $\omega^2$ can be characterized as the first limit element with more than finitely many preceding limit elements, viz. the elements $\omega, 2\omega, 3\omega, \ldots$. We continue and find many further elements: $\omega^2 + 1, \omega^2 + \omega, \omega^2 + 2\omega, \ldots$ until we come to $2\omega^2$ which follows immediately after the second $\omega^{2\omega}$-block which starts with $\omega^2$:

We continue with $\omega^{2\omega}$-blocks and after all the elements $0, \omega, \omega^2, \omega^3, \omega^4, \ldots$ comes the element we denote $\omega^3$. Then we go on with $\omega^{3\omega}$-blocks and after infinitely many of these reach out to $\omega^4$. Then with $\omega^{4\omega}$-blocks we get at $\omega^5$ and so on and so on.

After the elements $\omega, \omega^2, \omega^3, \ldots$ we still have only seen a countable part of $X$. The first element that remains is denoted $\omega^\omega$. After $\omega^\omega$ we find elements such as, $2\omega^\omega, \omega^\omega + 1, \omega^\omega + 2, \omega^{2\omega}, \omega^{3\omega}, \omega^{2\omega^2}, \omega^{3\omega^2}, \omega^{2\omega^3}, \omega^{3\omega^3}, \omega^{2\omega^4}, \omega^{3\omega^4}, \ldots$ with many many, in fact infinitely many omissions. After $\omega, \omega^\omega, \omega^{2\omega}, \omega^{3\omega}, \ldots$ we run out of, so to speak, usual algebraic notation – but we do not run out of the set $X$. In spite of all our efforts we have still only seen a countable part of $X$.

The first element not yet seen, after having run out of normal algebraic notation we denote $\varepsilon_0$. This is the first $\varepsilon$-ordinal number. And we could continue $(2\varepsilon_0, \varepsilon_0^2, \varepsilon_1, \varepsilon_2, \varepsilon_\omega, \varepsilon_{\varepsilon_0}, \ldots)$, but now we find that enough is enough and pause.

What we have learned is that with our descriptive power we cannot constructively encompass more than countably many elements of $X$. In this connection note that, accepting the set of natural numbers, all in
The above considerations was constructive in the sense that we could, if we were forced to do so, find models of all well-orderings met on the way. For instance, we can model what we met until \( \omega \), i.e. the left section \( V(\omega) \), which can of course be modelled by \( \mathbb{N} \) itself, and we can use copies of \( \mathbb{N} \) as building blocks to construct models of \( V(\omega^2) \) or more complicated parts of the well-ordering given on \( X \).

To go further we must leave the constructive approach and confine ourselves to descriptions. Recall that we assumed that the well ordering has a largest element, say \( x^* \). Then, as \( X \) is uncountable, the left section \( V(x^*) \) is uncountable. Using the property of the well ordering, we know that there exists a first element with this property. This element we denote \( \omega_1 \). Then we are this far:

\[
\begin{array}{cccccccccccc}
| & | & \cdots & | & \cdots & | & \cdots & | & \cdots & | & \cdots & | & \cdots & | & \cdots & | & \cdots & | & \cdots & | & \cdots & | & \cdots & |
\end{array}
\]

\[
0
\begin{array}{cccccccccccc}
\omega & \omega^2 & \omega^3 & \omega^\omega & \omega^{\omega^\omega} & \varepsilon_0 & \omega_1
\end{array}
\]

The new ordinal number \( \omega_1 \) is characterized by the following properties:

- \( V(\omega_1) \) is uncountable,
- For all \( \beta < \omega_1 \), \( V(\beta) \) is countable.

So \( \omega_1 \) is exactly where we got out of the “countability jungle”. Finally we reached out to the uncountable, finally we ran out of Hilbert’s Hotel. So the set \([0, \omega_1]\) of all ordinal numbers on the way to \( \omega_1 \) cannot be accommodated in Hilbert’s Hotel – but can easily be accommodated in Jessen’s ballroom, say as the set of boys.

On our way from the start, 0, to \( \omega_1 \) we always met something new, a new well ordered set, the left section associated with the current step in the construction. But whereas we met sets with more and more elements out to \( \omega \), in a certain sense nothing happens on the long way from \( \omega \) until (but not including) \( \omega_1 \) as all the sets we met there could be accommodated in Hilbert’s Hotel. But, paying also attention to the order in which the elements occurred, something did happen, the sets got more and more complicated. And finally, when we are at the ordinal \( \omega_1 \) the miracle happens, suddenly we are faced with a truly larger set, an uncountable set. And we realize that though in the sense of equipotence nothing happened between \( \omega \) and \( \omega_1 \), all the intervening steps were necessary in order to reach out that far, to our first uncountable set. The situation is similar to the “miracle” that happened at \( \omega \): The left section \( V(\omega) \) is infinite, whereas all preceeding left sections are finite.
The human brain finds it difficult to comprehend an element as $\omega_1$. Anyhow this and similar elements are not that difficult to work with. The truly disturbing jungle of ordinal numbers needed to reach out to $\omega_1$ can in many contexts be forgotten or we just use the fact that these numbers are there without having to worry about their, admittedly immensely complicated structure.

### 3.7 Ordinal and cardinal numbers

We continue our study of the “huge” set $X$ from the previous section. Now we pay main attention to the steps in our construction where “miracles” happen, i.e. where larger sets, in the sense of equipotence, emerge. We only look at infinite sets. The “miracle” ordinal numbers are called **cardinal numbers**. These sets are given special names, they are the **alephs**, denoted $\aleph_0$, $\aleph_1$ and so on. Thus, some of the numbers we deal with have double names, one when conceived as ordinal numbers, another when conceived as cardinal numbers. The first double name is attached to $\omega$ which is also denoted $\aleph_0$. It is the first infinite cardinal number according to the following definition:

**Definition 3.7.1.** A **cardinal number** is an ordinal number $\beta$ such that no left section $V(\alpha)$ with $\alpha < \beta$ is equipotent with $V(\beta)$. A cardinal number is **finite** if its left section is finite. Otherwise it is **infinite** or it is said to be an **aleph**.

So $\aleph_0$ is the first aleph. The next aleph is $\omega_1$, which is also denoted $\aleph_1$.

The definition presupposes that we continue our construction of ordinal numbers. So far we got to $\omega_1$. But, as $X$ is huge, we can continue the construction. For a lot of the ordinal numbers constructed after $\omega_1$ – with $\omega_1 + 1$ being the first one – we do not see an increase in equipotence. But then, at a certain ordinal number which we denote $\omega_2$ we do see an increase. In more detail, $\omega_2$ is defined using the well ordering property as the first $x \in X$ for which $|V(x)| > |V(\omega_1)|$. That such an element exists is clear as $|V(x^+)| = |X| > |V(\omega_1)|$ since $X$ is huge. We realize that $\omega_2$ is characterized by the conditions:

(i) $|V(\omega_2)| > |V(\omega_1)|$,

(ii) for all $\alpha < \omega_2$: $|V(\alpha)| \leq |V(\omega_1)|$.

Clearly, $\omega_2$ is the next aleph after $\aleph_1$. It is denoted $\aleph_2$.

Once more looking at our construction of ordinal numbers we can
The infinite

proceed well beyond \( \omega_2 \) if only \( X \) is large enough. So, only sketching the position of the alephs, we get at the numbers:

\[
\begin{array}{cccccccccccc}
| & \cdots & | \cdots | & \cdots & | \cdots | & \cdots & | \cdots | & \cdots & | \cdots |
\end{array}
\]

\[
\begin{array}{cccccccccccc}
\aleph_0 & \aleph_1 & \aleph_2 & \aleph_3 & \aleph_\omega & \aleph_{\omega+1} & \aleph_{\omega+2} & \aleph_{2\omega} & \cdots
\end{array}
\]

We stress that though we have left out the ordinal numbers which are not alephs (except for the very start, 0) these ordinal numbers are essential. Without them we could never reach out to the alephs.

\[
\begin{array}{cccccccccccc}
| & \cdots & | \cdots |
\end{array}
\]

\[
\begin{array}{cccccccccccc}
\aleph_\omega & \aleph_{\omega+1} & \aleph_{\omega+2} & \cdots
\end{array}
\]

Hertil nede vi med \( X \)

\[
\begin{array}{cccccccccccc}
| & \cdots & | \cdots & | \cdots |
\end{array}
\]

\[
\begin{array}{cccccccccccc}
\aleph_\omega & \aleph_{\omega+1} & \aleph_{\omega+2} & \cdots
\end{array}
\]

og hertil med \( X! \)

3.8 Exercises

(i) You may have met the infinite in other ways than indicated in Section 3.1. And surely, you have heard of paradoxes such as Archimedes and the turtle. Discuss! (Possibilities: Identify other encounters with the infinite, discuss paradoxes about the infinite and how they can be illuminated by present day modelling allowing for infinite objects or infinite processes).

(ii) Which natural number did Archimedes find as an upper bound in his count of sand grains? (Search after Archimedes, the sand reckoner, King Gelon or ...).

(iii) Consider other infinite objects which exist according to the (uncritically accepted) idea of comprehension. Do you have difficulties using the idea to reach acceptance of the set of real numbers? The issue is discussed later in Section xx.x.

(iv) Consider the set \( \mathbb{Z} \) of integers as well as the set of lattice points in the plane (points of the form \((n,m)\) with \(n\) and \(m\) integers). Discuss how these sets can be formed as well-defined sets, not by using comprehension, but using simple constructs applied to the set \( \mathbb{N} \) which is assumed to be well defined. See further in Exercise xx.x.

(v) The fascination of Hilbert’s Hotel has been rather overwhelming. Enter “Hilbert’s Hotel” on a search engine such as Google and see what you find. Collect the material appropriately and present what you find interesting in class!
(vi) Show that there are no more rational numbers than natural numbers in the sense that these numbers can all be assigned a room number in Hilbert’s Hotel - without duplication, of course. Hint: Use the idea from the third story about Hilbert’s hotel, noting that every rational number has a numerator and a denominator.

(vii) The goal here is to show that there are no more eight-tuples of natural numbers than natural numbers, understood in the usual way. This can be done in many ways. One idea is to break eight-tuples up in pairs, another is to accommodate eight tuples according to their “trace” (i.e. the sum of the entries). Use one of these methods – or any other method of your own choice – to find a formula or an algorithm which can be used to assign a room number to every eight-tuple of natural numbers. If you wish, you may comment on the complexity of the procedure suggested, understood in vague terms as adequate comments, possibly of a quantitative nature which illuminates how complicated your method is.

Use your method to calculate the room number of inhabitant \((1,2,3,4,5,6,7,8)\) in the galaxi version of stories about Hilbert’s hotel.

(viii) Let \(N \in \mathbb{N}\). Show that all \(N\)-tuples of natural numbers can be accommodated in Hilbert’s Hotel.

(ix) It may seem unnecessary to have the special Theorem 3.4.7 first and then the general Theorem 3.4.8. And indeed it is. Modify the formulation of Theorem 3.4.7 so that it appears as a special case of Theorem 3.4.8.

(x) Note that Theorem 3.4.8 holds for any set, finite or not. What does the result say in case \(X = \emptyset\)?

(xi) Suggest a way to define a sequence \((X_0, X_1, X_2, \cdots)\) of sets representing different infinities \((|X_m| \neq |X_n| \text{ for } m \neq n)\). Show that no matter how you set-up such a sequence, you can always find a set representing a greater infinity than any of the sets in the sequence (there exists \(X \text{ with } |X| > |X_n| \text{ for all } n \geq 0\)). So there are more infinities than you can accommodate in Hilbert’s Hotel!

(xii) (i) Prove that if \(Y_0\) is a subset of a set \(Y\) and if \(|Y_0| > |N|\), then \(|Y| > |N|\).

(ii) Prove that \(|\mathbb{R}| > |\mathbb{N}|\). Hint: Consider the subset of \(\mathbb{R}\) consisting of all \(t \in [0,1]\) which can be represented as a decimal fraction using only 0’s and 1’s and reason as in the proof of Theorem 3.4.7.
The infinite

(iii) Give an alternative proof of the relation $|\mathbb{R}| > |\mathbb{N}|$ by exploiting the fact that if $F_1 \supseteq F_2 \supseteq F_3 \supseteq \cdots$ is a sequence of non-empty closed subsets of $\mathbb{R}$ whose diameters converge to 0, then the sequence contains a common point. Hint: Consider a numbering of (some) real numbers in $[0, 1]$ and construct smaller and smaller intervals such that the $n$’th interval does not contain the $n$’th real number.

(xiii) Prove that $[0, 1]$ and $[0, 1]$ are equipotent.

(xiv) Using Theorem ?? prove that if $X$ is infinite and $x_0 \in X$, then $X$ and $X \setminus \{x_0\}$ are equipotent. Generalize the result, taking away more elements from $X$.

3.9 Some further exercises

Note: The exercises below are largely adapted from exercises in my MatY-notes (in danish). There may be a little overlap with previous material. Some of the exercises depend on ZFC and on certain consequences, e.g. that every set can be well ordered (Zermelo’s well ordering theorem) and also some results on cardinal numbers and cardinal arithmetic, though you are asked to prove parts of this in the exercises. Only little care has been taken to arrange the exercises in a logical order.

(i) Prove that a countable union of countable sets is countable.

(ii) An indicator function is a function with range contained in $\{0, 1\}$. Let $X$ be a set. Show that there is a natural bijection between the power set of $X$ and the set of indicator functions defined on $X$, hence the two sets are equipotent.

(iii) If $X$ is infinite and $|X| < |Y|$, then, for every injection $f : X \hookrightarrow Y$, $|B| = |Y|$ with $B$ the set of “wallflowers” ($B = Y \setminus f(X)$). Hint: Perhaps this is not so easy, but you can prove this directly. You may also choose to use the result of a later exercise involving sums of cardinal numbers.

(iv) (i) For every sequence of sets $(X_n)_{n \geq 1}$ there exists a set $X$ with $|X| > |X_n|$ for every $n \geq 1$.

(ii) For every sequence of cardinal numbers, there is a cardinal number larger than any of the cardinal numbers in the sequence.

(iii) Is it also true that for any set of cardinal numbers, there is a cardinal number larger than any of the cardinal numbers in the set?
(v) Let $A$ denote the set of *algebraic numbers*, real numbers which are roots in a polynomial with integer coefficients. Prove that the sets $N, Z, Q, A$ and $N^{(N)}$ are all equipotent (regarding the last set, see the definition in Exercise 2.5).

(vi) Define a *cardinal number* as an ordinal number $\kappa$ which is the first in its equipotence class in the sense that $|\alpha| < |\kappa|$ for every ordinal number $\alpha$ with $\alpha < \kappa$. Note that every natural number is a cardinal number and show that the infinite cardinal numbers can be given unique names $\aleph_\alpha$ with the $\alpha$’s ranging over the class of all ordinal numbers in such a way that

$\aleph_0 = \omega,$

$\aleph_{\alpha + 1}$ is the first cardinal number larger than $\aleph_\alpha$ for every ordinal $\alpha$,

$\aleph_\gamma = \sup_{\beta < \gamma} \aleph_\beta$ for every limit ordinal $\gamma$.

(vii) Let $\kappa$ and $\lambda$ be cardinal numbers. Choose sets $X$ and $Y$ such that $\kappa = |X|$, $\lambda = |Y|$ and (because of the definition below of the sum) $X \cap Y = \emptyset$. Then define

\[
\kappa + \lambda = |X \cup Y|,
\]

\[
\kappa \cdot \lambda = |X \times Y|,
\]

\[
\kappa^\lambda = |X^Y|.
\]

Show that these definitions make sense and prove (at least some of) the following (elementary!) results:

(i) $\kappa + \lambda = \lambda + \kappa$,

(ii) $\kappa + (\lambda + \mu) = (\kappa + \lambda) + \mu$,

(iii) $\kappa_1 \leq \lambda_1 \land \kappa_2 \leq \lambda_2 \Rightarrow \kappa_1 + \kappa_2 \leq \lambda_1 + \lambda_2 \land \kappa_1 \cdot \kappa_2 \leq \lambda_1 \cdot \lambda_2 \land \kappa_1^2 \leq \lambda_1^2$,

(iv) $\kappa \cdot \lambda = \lambda \cdot \kappa$,

(v) $\kappa \cdot (\lambda + \mu) = \kappa \cdot \lambda + \kappa \cdot \mu$,

(vi) $\lambda > 0 \Rightarrow \kappa \cdot \lambda \geq \kappa$,

(vii) $\kappa + \kappa = 2 \cdot \kappa$,

(viii) $\kappa \cdot \kappa = \kappa^2$,

(ix) $\kappa^{\lambda + \mu} = \kappa^\lambda \cdot \kappa^\mu$,

(x) $(\kappa^\lambda)^\mu = \kappa^{\lambda \cdot \mu}$,

(xi) $(\kappa \cdot \lambda)^\mu = \kappa^\mu \cdot \lambda^\mu$. 

(viii) If \((X, \leq)\) is a well ordered set and \(\kappa\) a cardinal number such that every left section of \(X\) has cardinality strictly less than \(\kappa\), then \(|X| \leq \kappa\). Does a stronger result hold where “strictly less than” is replaced by “\(\leq\)”? 

(ix) Let \(\kappa\) and \(\lambda\) be cardinal numbers with \(\kappa \leq \lambda\) and \(\lambda\) infinite. Show that then 

1. \(\kappa + \lambda = \lambda\). 
2. \(\kappa \geq 1 \Rightarrow \kappa \cdot \lambda = \lambda\). 
3. \(\kappa \geq 2 \Rightarrow \kappa^\lambda = 2^\lambda\). 

Hints: Once you have proved that for all infinite cardinal numbers \(\kappa\), 

\[(*) \quad \kappa \cdot \kappa = \kappa,\]

the rest follows without much difficulty (indication: (1): \(\lambda \leq \kappa + \lambda \leq \lambda \cdot \lambda = \lambda\). (2): \(\lambda \leq 1 \cdot \lambda \leq \kappa \cdot \lambda \leq \lambda \cdot \lambda = \lambda\). (3): \(2^\lambda \leq \kappa^\lambda \leq \lambda^\lambda \leq (2^\lambda)^\lambda = 2^{\lambda \times \lambda} = 2^\lambda\)).

To prove the central property \((*)\) use induction. Assume for \(\kappa\) an infinite cardinal number that for every cardinal number \(\alpha < \kappa\), either \(\alpha\) is finite or else \(\alpha \cdot \alpha = \alpha\). We must prove that \(\kappa \cdot \kappa = \kappa\). 

To do so we find a wellordering on \(\kappa \times \kappa\), such that \(|V(\xi)| < \kappa\) for every left section \(V(\xi)\). The desired wellordering is given by first ordering “after the diagonal, then lexicographically” in more detail, \((a, b) \leq (a', b')\) if and only if \((\max(a, b) < \max(a', b'))\) or \((\max(a, b) = \max(a', b')\) and \((a, b) \leq (a', b')\), lexicographically. This is indeed a wellordering. If \(\xi = (a, b) \in \kappa \times \kappa\), then \(V(\xi) \subseteq \alpha \times \alpha\), where \(\alpha = \max(a, b)\). 

If \(\alpha\) is finite, \(V(\xi)\) is finite and \(|V(\xi)| < \kappa\). And if \(\alpha\) is infinite, \(|V(\xi)| \leq |\alpha \times \alpha| = \alpha < \kappa\). Then use the result of the previous exercise.

(x) To prove: There exists an ordinal number \(\alpha\) such that \(\aleph_\alpha = \alpha\). 

Hint: Construct cardinal numbers \(\sigma_0, \sigma_1, \ldots\) by \(\sigma_0 = \aleph_0, \sigma_n = \aleph_{\sigma_{n-1}}\). Put \(\alpha = \sup_{n \geq 0} \sigma_n\). 

Remarks: \(\alpha = \aleph_\alpha\) looks large. It need not be that large, in fact, it is consistent with ZFC that \(2^{\aleph_\alpha} > \alpha\) (but \(2^{\aleph_\alpha} = \alpha\) is not possible!)
3.9 Some further exercises

(xi) Let $\kappa$ be a cardinal number. A set $A \subseteq [0, \kappa]$ is **cofinal** in $\kappa$ if $A$ or “unbounded”, i.e. $\forall \beta < \kappa \exists \alpha : \alpha \in A \land \alpha \geq \beta$. The **cofinality** $\text{cf}(\kappa)$ of $\kappa$ is the smallest cardinality of a set which is cofinal in $\kappa$. We call $\kappa$ **regular**, if $\text{cf}(\kappa) = \kappa$.

To prove: The first regular cardinal number larger than 1 is $\aleph_0$. Also prove that $\aleph_1$ is regular whereas $\aleph_\omega$ is not. Given the energy, also prove that $\aleph_{\alpha + 1}$ is regular for every ordinal number $\alpha$.

(xii) Prove that the sets

$$2^\mathbb{N}, \ \mathbb{N}^\mathbb{N}, \ (2^\mathbb{N})^\mathbb{N} \text{ and } (\mathbb{N}^\mathbb{N})^\mathbb{N}$$

are equipotent.

(xiii) Prove that the sets

$$\mathbb{R} \setminus \mathbb{Q}, \ \mathbb{R}, \ \mathbb{R}^2, \ \mathbb{R}^3, \ldots \text{ and } \mathbb{R}^\mathbb{N}$$

are all equipotent with $2^\mathbb{N}$.

(xiv) Show that $\mathbb{R}^2$ and $2^\mathbb{R}$ are equipotent and conclude that there are exactly as many functions $f : \mathbb{R} \to \mathbb{R}$ as there are subsets of $\mathbb{R}$. Thus, there are more real functions of a real variable than there are real numbers.

(xv) Generalize the previous result and show that for any infinite set $X$, $X^X$ and $2^X$ are equipotent.

(xvi) Call an interval $I \subseteq \mathbb{R}$ **rational**, if both endpoints are rational and call a rectangle $I \times J \subseteq \mathbb{R}^2$ **rational**, if both $I$ and $J$ are rational. Show that the set of rational rectangles in $\mathbb{R}^2$ is countable.

Show that there exists a sequence of rational rectangles, even squares, whose union is the open unit disk.

Call (or recall that) a subset $G \subseteq \mathbb{R}^2$ **open**, if, for every $(x, y) \in G$ there exists $\varepsilon > 0$, such that the disk with centre $(x, y)$ and radius $\varepsilon$ is contained in $G$. Show that every open subset of $\mathbb{R}^2$ can be represented as a countable union of rational rectangles. Conclude from this that the set of all open subsets of $\mathbb{R}^2$ is equipotent with $\mathbb{R}$ itself. Thus, in a sense, there are very few such subsets.

(xvii) Prove that

$$|C(\mathbb{R})| = |\mathbb{R}| = |2^\mathbb{N}|,$$

where $C(\mathbb{R})$ denotes the set of continuous real functions defined on $\mathbb{R}$. Thus, in a sense, there are very few such functions.

*Hint:* Identify a continuous function with its graph and look at the complement of the graph and exploit the result of exercise...
(xvi). Another, arguably more natural, idea is to exploit the fact that a continuous function may be identified from its values on a dense set (such as \( \mathbb{Q} \)).

(xviii) Let \( M \) denote the set of monotonically increasing functions \( \mathbb{R} \to \mathbb{R} \). Show that there are, in a sense (see previous exercises), very few such functions.

   \textit{Hint:} A monotone function may be identified by its discontinuity points, its values at these points and its values at the rational points. (By the way, do construct an increasing function with \( \mathbb{Q} \) as its set of discontinuity points!)

(xix) How “many” polygons are there in \( \mathbb{R}^2 \), when we also count the edges as belonging to the polygon? Does your answer change if we may count only parts of the edges as belonging to a polygon?

(xx) Prove that for every infinite set \( X \), the set of bijections \( X \to X \) is equipotent with \( 2^X \). \textit{Note:} I am not sure how to do this in the most elegant or elementary way. The proof I have in mind uses the fact (not difficult to prove!) that for every set which is not empty or a singleton, there exists a bijection of the set to itself which has no fixpoints (points with \( f(x) = x \)). For \( X = \mathbb{R} \) the result is easy to prove if you assume that the continuum hypothesis holds. In any case there are “incredibly many” bijections from \( \mathbb{R} \) to \( \mathbb{R} \) but “hardly any” continuous or monotone real valued functions on \( \mathbb{R} \).

(xxi) Sierpinski showed that there exists a family \((A_i)_{i \in I}\) of subsets of \( \mathbb{N} \) such that

1) \( \forall i : A_i \) is infinite,

2) \( \forall i \neq j : A_i \cap A_j \) is finite,

3) \( I \) is equipotent with \( 2^{\mathbb{N}} \).

Prove this result. \textit{Hint:} Every real number is the limit of a sequence of rational numbers!

   Can you replace the second condition with the requirement that the \( A_i \)'s are pairwise disjoint?
4

Axiomatic set theory

This chapter is somewhat telegraphic in style. You may read more in my MatY notes in Danish or in Hrbacek & Jech: Introduction to Set Theory, Marcel Dekker, 3rd edition 1999.

4.1 Why axiomatize?

You may aim for

(i) consistency!
(ii) rich basis for all mathematics!
(iii) only “obviously true” axioms!
(iv) complete freedom!
(v) making development of mathematics easier!
(vi) develop aestethically attractive, beautiful theory!
(vii) completeness!
(viii) completely safe basis for your reasoning!

Without going into a full discussion, the table below summarizes what can be achieved.
4.2 ZFC—a survey

Axiom 0: (Set existence). There exists a set \((\exists x : x = x)\).

Axiom 1: (Extensionality). A set is determined by its elements. Formally: \(\forall x \forall y : (\forall z : z \in x \iff z \in y) \Rightarrow x = y\).

Axiom 2: † (The axiom of foundation or the regularity axiom). If \(x \neq \emptyset\), there is \(y \in x\) such that \(x \cap y = \emptyset\), i.e. for all \(z \in x\), \(z \notin y\) holds.

Axiom 3: ‡ (Comprehension axiom). For every predicate \(P(x)\) with \(x\) as free variable and every set \(A\), there is a set \(y\) consisting of all elements \(x \in A\) which have the property \(P\). Thus: \(x \in Y \iff x \in A \land P(x)\). Notation: \(\{x \in A | P(x)\}\).

Axiom 4: (The axiom of (unordered) pairs). For all \(x\) and \(y\) there is a set whose elements are \(x\) and \(y\). Formally:

\[ \forall x \forall y \exists z \forall w : w \in z \iff w = x \lor w = y. \]

Notation: \(\{x, y\}\).

Axiom 5: § (The sum axiom). For every set \(A\) the union of sets in \(A\) is a set. Formally:

\[ \forall A \exists y \forall (y \in x \iff \exists a \in A : y \in a). \]

Notation: \(\cup A\) eller \(\bigcup_{a \in A} a\).

† A convenient technical axiom which, in a certain sense, is not necessary.

‡ Det giver normalt ikke problemer, at anvende dette vigtige aksiom, men strengt taget bør det præciseres, hvilke åbne udsagn, det drejer sig om. De grundlæggende udsagn er \(x \in y\) og \(x = y\). På dem kan de logiske konnektiver \(\neg\) (non), \(\land\) (konjunktion) og \(\lor\) (disjunktion) anvendes, samt logisk kvantivering v.hj. af \(\forall\) (alkvantoren) og \(\exists\) (eksistenskvantoren). Herved fremkommer åbne udsagn med en række variable. I aksiomet skal det forudsettes, at \(x\) er en fri variabel, dvs. ikke bundet til alkvantorer eller eksistenskvantorer. Bemærk, at aksiomet er et aksiomsskema med uendeligt mange aksiomer, ét for hvert åbent udsagn af den betragtede type.

§ We may write \(\cup A = \{y \mid P(y)\}\), where \(P(y)\) is the predicate \(\exists a \in A : y \in a\). Note that the existence of this set does not follow from the comprehension axiom.
Axiom 6: *(The substitution axiom).* Let $P(x,y)$ be a predicate in $x$ and $y$, such that, to any $x$ there exists at most one $y$, such that $P(x,y)$ holds. Then, to every set $A$, a set $B$ exists, such that $y \in B \iff \exists x \in A : P(x,y)$.

Axiom 7: *(The axiom of infinity).* There exists an inductive set, i.e. a set $x$ such that $0 \in x$ and such that $y \in x \Rightarrow y \cup \{y\} \in x$ (here, $0 = \emptyset$).

Axiom 8: *(The power set axiom).* For every set $x$ the set of all subsets of $x$ exists, i.e., denoting this set by $\mathcal{P}(x)$, for all $y, y \in \mathcal{P}(x) \iff y \subseteq x$ holds.

Axiom 9: *(The axiom of choice (AC)).* For every set $A$, whose elements are non-empty sets there exists a choice function, i.e. a function $\varphi$ defined on $A$ such that $\varphi(x) \in x$ for all $x \in A$.

4.3 Elementary constructions

The reader is supposed to be familiar with basic constructs of ZFC such as successor ($Sx = x \cup \{x\}$), small natural numbers or $0$ ($0 = \emptyset, 1 = S0, 2 = S1$ etc.), ordered pairs, relations, special relations (equivalence relations, various types of order relations), functions (various notions related to the notion of a function), union, intersection, difference set, symmetric difference set, quotient set, product sets etc. This material can be found in danish in MatY and in english in HJ.

The main point is that all these notions, well known from naive set theory, can be defined rigorously inside ZFC and their key properties developed. And here the focus is on the definitions whereas the development of key properties follows the development known from naive set theory so closely that there is no need to repeat all the well known arguments. All that is needed is a little care so that you ensure that you always work inside ZFC (e.g. that you do not work with the class of all sets as if this is a set) and then it is good practice to note every time you use the only non definite or non constructive axiom of ZFC, the axiom of choice, AC. Here we do not have sophisticated applications of AC in mind – this we return to in Section xx – but simple naive constructions such as the right inverse of a surjective function.

Perhaps I should remind you specifically of the notion of disjoint sum (or union) as this may not be so well known. The idea is to consider the usual union but paying also attention to which set an element in the union is considered to come from. Formally,

$$\uplus x = \{(y,z) | y \in x, \ z \in y\}.$$
As usual uniqueness is ensured by extentionality. Existence is ensured by comprehension, indeed, the set is a subset of \( \times \cup x \). If we have given an indexed family of sets, say \((A_i)_{i \in I}\), we may use instead a variant of the construction and consider the set of all pairs \((i, z)\) with \(i \in I\) and \(z \in A_i\).

In my preferred type of presentation of the above indicated material, I use a treasure box which is nothing but a model of ZFC, conceived as a box containing all the “pieces” the mathematician is allowed to play with, much as a box of LEGO pieces. This is a pedagogic device in line with Hilbert’s Hotel and Jessen’s Ballroom, which may be especially useful if you only know little of logic and model theory. For details see my MatY-notes.

### 4.4 Ordinal numbers

We embark on the not-so-elementary (or naive) constructions based on ZFC. These are also where the main profit of formalization, axiomatization lies in that it is here we

- lay the ground for the precise definitions of basic concrete sets as the natural numbers, the reals etc.
- develop the basis needed in order to deal rigorously with the notion of infinity
- establish a basic general tool, constructions by recursion, which is a key to many constructions of present day mathematics and computer science.

The key to all this is the introduction of ordinal numbers – done in MatY first at the naive level – and the development of their basic properties. Though the precise development is available in danish in MatY, being of central importance for the course, I decided to prepare a translation, actually a brief translation taken directly from MatY. The material is also in HJ, but with some variations.

Technically we make extensive use of the axiom of foundation:

\[
\forall x \neq 0 \exists y \in x \exists z \in x : z \notin y.
\]

A set \(y\) with the indicated property \((y \in x \land y \cap x = \emptyset)\) is called a foundation of \(x\). So every non-empty set has a foundation. Applying this to \(\{x\}\), we see that for every set \(x\),

\[
x \notin x
\]
Using foundation, allows us to give a relatively simple definition of ordinal numbers. First we note some simple consequences of the axiom. In order to avoid repetition, we state these consequences as if the set of natural numbers is already known. Thus the results are to be proved at the naive level and only get their full strength once the set of natural numbers has been properly defined. The reader must watch out that only instances of the results—such as the above stated fact $\forall x : x \notin x$—which can be proved presently are used on the way to the definition of ordinals and natural numbers, to be dealt with in the sections to follow.

**Theorem 4.4.1.** The following holds:

(i) For every natural number $n$ and every finite sequence $x_1, \ldots, x_n$ of sets, there exists $i_0 \in \{1, \ldots, n\}$ such that, for all $i \in \{1, \ldots, n\}$, $x_i \notin x_{i_0}$.

(ii) There is no circular “membership-chain”, i.e. we cannot have $x_1 \in x_2 \in \cdots \in x_n$ with $x_1 = x_n$.

(iii) There is no infinite downwards filtering “membership-chain”, i.e. $x_1 \ni x_2 \ni x_3 \ni \cdots$ is impossible.

*Proof.* (i): Choose $x_{i_0}$ as a foundation of $\{x_1, x_2, \ldots, x_n\}$.

(ii):—If so, $\{x_1, x_2, \ldots, x_n\} = \{x_2, x_3, \ldots, x_n\}$ would not have a foundation.

(iii):—If so, $\{x_1, x_2, \ldots\}$ would not have a foundation. \(\square\)

**Definition** An ordinal or ordinal number is a set $\alpha$ such that:

(i) The membership relation is total on $\alpha$, i.e. $\forall x \forall y (x \in \alpha \land y \in \alpha \Rightarrow x \in y \lor y \in x \lor x = y)$.

(ii) $\alpha$ is transitive, i.e. $\forall x \forall y (y \in x \land x \in \alpha \Rightarrow y \in \alpha)$.

Transitivity says: $x \in \alpha \Rightarrow x \subseteq \alpha$.

**Theorem 4.4.2** (Basic properties of ordinals). *Der glder:*

(i) $\alpha$ ordinal number, $\beta \in \alpha \Rightarrow \beta$ ordinal number.

(ii) $\alpha$ ordinal number $\Rightarrow \alpha \cup \{\alpha\}$ ordinal number.

(iii) A union of ordinal number $\Rightarrow \cup A$ ordinal number.

(iv) A non-empty set of ordinal numbers $\Rightarrow \cap A$ ordinal number.

(v) For $\alpha$ and $\beta$ ordinal numbers, $\alpha \subseteq \beta \leftarrow \alpha \in \beta \lor \alpha = \beta$ holds.

(vi) For $\alpha$ and $\beta$ ordinal numbers, precisely one of the statements: $\alpha = \beta, \alpha \in \beta, \beta \in \alpha$ holds.
(vii) Let $\alpha$ be an ordinal number. Define, for $x \in \alpha$ and $y \in \alpha$, $x \leq y$ to mean $x \subseteq y$. Then “$\leq$” is a well ordering on $\alpha$. In this ordering, $x < y \iff x \in y$ and, for every $x \in \alpha$, $V(x) = x$ (recall the definition of the left section: $V(x) \overset{\text{def}}{=} \{y | y < x\}$).

(viii) A set $A$ of ordinal numbers such that $x \in a$ for an element $a \in A$ implies $x \in A$ is itself an ordinal number.

There is no set containing all ordinal numbers (we did this as an exercise). Anyhow, we introduce the notation $\text{ON}$ for the collection (class) of all ordinal numbers. We may apply notation such as $x \in \text{ON}$ and $A \subseteq \text{ON}$ as abbreviations for, respectively, “$x$ is an ordinal number” and “$A$ is a set and every element in $A$ is an ordinal number”.

**Proof.** (ii) is an exercise. To prove remaining properties, we shall establish the following properties:

(a) $\alpha \in \text{ON} \land \beta \in \alpha \Rightarrow \beta \in \text{ON}$.
(b) $A \subseteq \text{ON} \land A \neq \emptyset \Rightarrow \cap A \in \text{ON}$.
(c) $\alpha \in \text{ON} \land \beta \in \alpha \Rightarrow (\beta \subseteq \alpha \leftrightarrow \beta \in \alpha \lor \beta = \alpha)$.
(d) $\alpha \in \text{ON} \Rightarrow (\alpha, \leq)$ well ordering.
(e) $\alpha \in \text{ON} \land \beta \in \text{ON} \Rightarrow$ precisely one of: $\alpha = \beta, \alpha \in \beta$ and $\beta \in \alpha$ is true.
(f) $A \subseteq \text{ON} \Rightarrow \cup A \in \text{ON}$ (recall: When we write $A \subseteq \text{ON}$ it is assumed that $A$ is a set!).

(a): Let $x \in \beta \land y \in \beta \land x \neq y$. As $\alpha$ is transitive, $x \in \alpha \land y \in \alpha$. As “$\in$” is total on $\alpha$ and $x \neq y$, either $x \in y$ or $y \in x$ holds. Thus “$\in$” is total on $\beta$. To prove transitivity, assume $x \in y \in \beta$ (stenography for $x \in y \land y \in \beta$). Among the alternatives $x = \beta, \beta \in x$ and $x \in \beta$ (only possibilities as “$\in$” is total on $\alpha$ and as $x \in \alpha$ and $\beta \in \alpha$), the two first can be excluded (note that the axiom of foundation is used here). Thus $x \in \beta$. So $\beta$ is transitive. All in all, $\beta \in \text{ON}$. □

(b): We have: $\cap A = \{x | \forall \alpha \in A : x \in \alpha \}$. Assume that $x \in \cap A \land y \in \cap A$. Choose $\alpha \in A (A \neq \emptyset)$. Then $x \in \alpha \land y \in \alpha$ and as $\alpha \in \text{ON}$, either $x = y$ or $x \in y$ or $y \in x$ must hold. Hence “$\in$” is total on $\cap A$. Transitivity of $\cap A$: Easy exercise! □

(c): “$\Leftarrow$” i “$\Rightarrow$” easy exercise! To prove the opposite implication, assume that $\alpha$ and $\beta$ are ordinal numbers and that $\beta \subseteq \alpha$. If $\beta = \alpha$, all is OK. Assume, therefore, that $\beta \neq \alpha$. Then $\beta$ is a proper subset of $\alpha$ and hence $\alpha \setminus \beta \neq \emptyset$. Let $x$ be a foundation of $\alpha \setminus \beta$. To finish the proof, we shall show that $\beta = x$ (since, as $x \in \alpha$, then $\beta \in \alpha$ will hold).
First we prove that $x \subseteq \beta$: If $y \in x \setminus \beta$, $y \in \alpha \setminus \beta$ would hold, which contradicts $x \cap (\alpha \setminus \beta) = \emptyset$. Therefore, $x \setminus \beta = \emptyset$ must hold, i.e. $x \subseteq \beta$.

Then we prove that $\beta \subseteq x$: Assume that $y \in \beta$. As $\beta \subseteq \alpha$, $y \in \alpha$, hence – as both $x$ and $y$ are elements in the ordinal number $\alpha$ – one of the three alternatives $y = x$, $x \in y$ and $y \in x$ must hold. The two first are easily excluded. Thus $y \in x$, hence the implication $y \in \beta \Rightarrow y \in x$ is proved, and, as desired, $\beta \subseteq x$ follows.

(d): According to definition, for $x \in \alpha$ and $y \in \alpha$, $x \leq y$ means that $x \subseteq y$. Clearly, “$\leq$” is reflexive, antisymmetric and transitive (in brief: $x \leq x$, $x \leq y \land y \leq x \Rightarrow x = y$ and $x \leq y \land y \leq z \Rightarrow x \leq z$)—in fact, this holds for “$\subseteq$”, generally. To prove that “$\leq$” is a well ordering of $\alpha$, consider a non-empty set $x \subseteq \alpha$. Let $y$ be a foundation of $x$. Then $y \in x$. In order to prove that $y$ is the first element in $x$, consider an arbitrary element $z \in x$. As $y \cap x = \emptyset$, $\gamma(z \in y)$ must hold. Then $(y \in z) \lor (y = z)$ follows. As $y$ as well as $z$ are ordinals (consider!), it follows from (c) (the easy part!) that $y \subseteq z$, i.e. $y \leq z$. Thus $\forall z (z \in x \Rightarrow y \leq z)$ as desired.

(e): Put $\gamma = \alpha \cap \beta$. Then $\gamma \in \Omega$. As $\gamma \subseteq \alpha$, it follows by (c) (the difficult part!), that $(\gamma \in \alpha) \lor (\gamma = \alpha)$. From $\gamma \subseteq \beta$ we conclude in a similar way that $(\gamma \in \beta) \lor (\gamma = \beta)$. It is not possible that both $\gamma \in \alpha$ and $\gamma \in \beta$ hold as then we would have $\gamma \in \alpha \cap \beta = \gamma$! We conclude that $(\gamma = \alpha) \lor (\gamma = \beta)$, i.e. that $(\alpha \subseteq \beta) \lor (\beta \subseteq \alpha)$. Apply (c) once more and conclude that $(\alpha \in \beta \lor \alpha = \beta) \lor (\beta \in \alpha \lor \beta = \alpha)$, in other words, $(\alpha = \beta) \lor (\alpha \in \beta) \lor (\beta \in \alpha)$. Clearly only one of these alternatives can hold.

(f): Let $A \subseteq \Omega$ and put $x = \cup A$. Then $x = \{y | \exists \alpha \in A : y \in \alpha\}$. To prove that “$\subseteq$” is total on $x$, assume $y \in x$ and $z \in x$. Determine $\alpha \in A$ and $\beta \in A$, such that $y \in \alpha$ and $z \in \beta$. By (e), either $\alpha \in \beta$, $\alpha \in \beta$ or $\beta \in \alpha$ holds. In all cases, there exists an ordinal number (either $\alpha$ or $\beta$) which contains both $z$ and $y$ as elements. As “$\subseteq$” is total on every ordinal number, we conclude that $(z = y) \lor (z \in y) \lor (y \in z)$, as desired. In order to prove that $x$ is transitive, assume that $y \in z \in x$. Determine $\alpha \in A$ such that $z \in \alpha$. Then $y \in z \in \alpha$, hence $y \in \alpha$. As $y \in \alpha \in A$ we conclude, as desired, that $y \in \cap A = x$.

The ordering introduced in the theorem is, strictly speaking, an ordering on $\alpha$ for $\alpha$ for a fixed ordinal number. However, we consider “$\subseteq$”
Axiomatic set theory

on all of \( \mathbb{ON} \), when only we recall that \( \mathbb{ON} \) is not a set. Thus \( (\mathbb{ON}, \leq) \) is not an ordered set. In spite of this it is convenient to talk about many concepts as if \( (\mathbb{ON}, \leq) \) is an ordering in the usual (strictly set theoretical) sense. Normally this does not give rise to any confusion. Examples illustrating the convenience of this somewhat sloppy practise is illustrated in the theorem below where we identify in order theoretical terms the ordinal numbers we found in (ii), (iii) and (iv) in the theorem above.

**Theorem 4.4.3.**

(i) Let \( \alpha \in \mathbb{ON} \). Then \( \alpha \cup \{\alpha\} \) is the least ordinal number, larger than \( \alpha \), the successor of \( \alpha \).

(ii) Let \( A \) be a set of ordinal numbers. Then \( \cup A = \sup A \), i.e. \( \cup A \) is the least ordinal number, which is \( \geq \) every ordinal number in the set \( A \).

(iii) Let \( A \) be a non-empty set of ordinal number. Then \( \cap A = \min A \), i.e. \( \cap A \) is the first ordinal number in \( A \).

**Proof.** (i): Put \( \beta = \alpha \cup \{\alpha\} \). Then \( \alpha \subseteq \beta \) and \( \alpha \neq \beta \), hence \( a < \beta \). Assume that also the ordinal number \( \gamma \) is larger than \( \alpha \): \( \alpha < \gamma \). Then \( \alpha \subseteq \gamma \) and \( \alpha \in \gamma \). As \( \alpha \in \gamma \), \( \{\alpha\} \subseteq \gamma \). Then \( \alpha \cup \{\alpha\} \subseteq \gamma \) must hold, i.e. \( \beta \subseteq \gamma \) or \( \beta \leq \gamma \), as desired.

(ii): Put \( \beta = \cup A \). Clearly, \( \forall \alpha \in A : \alpha \subseteq \beta \). Therefore, \( \forall \alpha \in A : \alpha \leq \beta \). Assume now that also the ordinal number \( \gamma \) has this property (\( \forall \alpha \in A : \alpha \leq \gamma \)). Then, \( \forall \alpha \in A : \alpha \leq \gamma \), hence \( \cup A \subseteq \gamma \), i.e. \( \beta \subseteq \gamma \) or \( \beta \leq \gamma \), as desired.

(iii): Let \( \alpha_0 \in A \) and put \( A_0 = \{\alpha \in A | \alpha \leq \alpha_0 \} = \{\alpha \in A | \alpha \subseteq \alpha_0 \} \). Then \( \cap A = \cap A_0 \). As \( A_0 \subseteq \alpha_0 \cup \{\alpha_0\} \), it follows from (i) and af (vii) in Theorem xx, that \( A_0 \) has a first element, say \( \gamma \). Now \( \cap A_0 = \gamma \) easily follows. \( \square \)

### 4.5 Induction and recursion

Ordinal Numbers are decomposed into three categories. The first only contains the one ordinal number 0. It is the first of all ordinal numbers. The next category contains all *successor ordinal numbers* (or just successors), i.e. ordinal numbers of the form \( \beta \cup \{\beta\} \) for a \( \beta \in \mathbb{ON} \). If \( \alpha \) is a successor, say \( \alpha = \beta \cup \{\beta\} \), we also write \( \alpha = \beta + 1 \). The last category contains all *limit ordinals*, all ordinal numbers which are distinct from 0 and not a successor.

Often we use “natural interval notation”, e.g. \([0, \alpha]\) for the set of ordinal numbers \( \beta \) with \( 0 \leq \beta < \alpha \) and \([0, \alpha]\) for \( \{\beta | 0 \leq \beta \leq \alpha\} \). Note that \([0, \alpha] = V(\alpha) = \alpha \) and \([0, \alpha] = [0, \alpha] \cup \{\alpha\} = \alpha \cup \{\alpha\} = \alpha + 1 \). When
we conceive $\alpha$ as a well ordered set, it is often better, more suggestive, to use the notation $[0, \alpha]$ instead of $\alpha$.

As $[0, \alpha]$ is a well ordered set, the usual \textit{induction principle} holds:

**Theorem 4.5.1** (Induction over $\alpha$). Let $[0, \alpha] = \alpha \in \mathbb{ON}$. If $x \subseteq [0, \alpha]$ satisfies the condition

$$\forall \beta < \alpha \ ([0, \beta] \subseteq x \Rightarrow \beta \in x),$$

then $x = [0, \alpha]$.

If we replace $(xx)$ by the conditions

$$0 \in x, \beta \in x \land \beta+1 < \alpha \Rightarrow \beta+1 \in x, \beta \in x \text{ for all } \beta < \gamma \land \gamma < \alpha \land \gamma \text{ limit ordinal} \Rightarrow \gamma \in x,$$

the same conclusion, $x = [0, \alpha]$ is obtained. (Exercise prove the theorem and the correctness of the remark just made).

From our naive discussion we know that we cannot always distinguish two ordinal numbers on ly viewing these as sets. But we can distinguish if we incorporate the order structure:

**Theorem 4.5.2.** Let $\alpha = [0, \alpha]$ and $\beta = [0, \beta]$ be ordinal numbers. If $[0, \alpha]$ and $[0, \beta]$ are order isomorphic, then $\alpha = \beta$ and the identity map is the only order isomorphism. \footnote{In various places, e.g. the MatY notes the notion of \textit{isomorphism} is studied. Let $(X, \leq)$ and $(Y, \leq)$ be ordered sets. The map $\phi : X \rightarrow Y$ is an \textit{order isomorphism} if 1) $\phi$ is bijective, if 2) $x_1 \leq x_2 \Rightarrow \phi(x_1) \leq \phi(x_2)$ and if 3) $y_1 \leq y_2 \Rightarrow \phi^{-1}(y_1) \leq \phi^{-1}(y_2)$. It is easily seen that 1) $\land$ 2) $\rightarrow$ 1) $\land$ 3) $\rightarrow$ 1) $\land$ 2) $\land$ 3).

\textit{Proof.} Assume that $\phi : [0, \alpha] \rightarrow [0, \beta]$ is an order isomorphism. Define $A \subseteq [0, \alpha]$ by $A = \{x \mid x < \alpha \land x < \beta \land \phi(x) = x\}$. By induction it is seen that $A = [0, \alpha]$. Then $[0, \alpha] \subseteq [0, \beta]$ must hold. Similarly, $[0, \beta] \subseteq [0, \alpha]$ must hold. Thus $[0, \alpha] = [0, \beta]$ (i.e. $\alpha = \beta$). From the induction result we see that $\phi$ is the identity on $[0, \alpha]$. \hfill \Box

Now turn to \textit{recursion}. By a \textit{transfinite sequence} we understand a map (syonymous for function) defined on an ordinal number. \footnote{though a bit misguiding, we even use this terminology when the ordinal number is finite} We often use “standard” sequence notation, $(f(x))_{x<\alpha}$ for $f$. We call $\alpha$ the \textit{length} of such a transfinite sequence.

A transfinite sequence with values in $X$ is a transfinite sequence $(f(x))_{x<\alpha}$ with $f(x) \in X$ for all $x < \alpha$, i.e., an element in $X^\alpha$. There is only one transfinite sequence of length $0$, viz. the empty sequence $0$. 

4.5 \textit{Induction and recursion}
Axiomatic set theory

(By convention, \(X^0 = \{\emptyset\}\).) By \(X^{(\alpha)}\) we denote the set of transfinite sequences of length less than \(\alpha\) and with values in \(X\):

\[X^{(\alpha)} = \bigcup_{\beta<\alpha} X^\beta\]

—this set is well defined as you will realize applying the axiom of substitution (look at \(\beta \ni X^\beta\) and apply also the sum axiom).

**Theorem 4.5.3** (Recursion over \(\alpha\)). Let \(\alpha\) be an ordinal number and \(Y\) a set. Let \(\Phi : Y^{(\alpha)} \to Y\) be a map. Then there exists a uniquely determined transfinite sequence \(f \in Y^{(\alpha)}\), such that

\[f(\beta) = \Phi( (f(x))_{x<\beta} ) \quad \text{for alle } \beta < \alpha.\]  

(4.1)

**Proof.** Uniqueness follows by induction over \(\alpha\).

Call \(g\) a partial solution if there exists \(\beta < \alpha\), such that \(g \in Y^\beta\) and \(g(\gamma) = \Phi( (g(x))_{x<\gamma} )\) for every \(\gamma < \beta\). If \(g \in Y^\beta\) and \(h \in Y^\gamma\) with \(\beta < \gamma\) are partial solutions, then, by induction, you see that \(h(x) = g(x)\) for all \(x < \beta\).

We then prove that there exist partial solution of any length \(\beta < \alpha\). For this, look at the set \(A = \{\beta < \alpha \mid \text{there is a partial solution } g \in Y^\beta\}\).

Let \(\beta < \alpha\) and assume, for the purpose of a proof by induction, that \([0, \beta) \subseteq A\). If \(\beta = 0\) or if \(\beta\) is a successor, it is easily shown that \(\beta \in A\).

Let us do it: If \(\beta = 0\), put \(g = \emptyset\). Then \(g \in Y^0\) with \(\beta = 0\). The condition \(g(\gamma) = \Phi( (g(x))_{x<\gamma} )\) for \(\gamma < \beta\) is empty in this case as there is no \(\gamma\) with \(\gamma < \beta\). And if \(\beta = \delta + 1\), define \(g \in Y^\beta\) by \(g(\gamma) = h(\gamma)\) for \(\gamma < \delta\), and \(g(\delta) = \Phi( (g(x))_{x<\delta} )\), where \(h \in Y^{\delta}\) is a partial solution.

Then \(g \in Y^\beta\) is a partial solution, so again \(\beta \in A\).

Then assume that \(\beta\) is a limit ordinal (still assuming \([0, \beta) \subseteq A\)). We now consider the predicate \(P(x, y)\) in two variables which is true precisely when:

either \((x \text{ is not an ordinal number less than } \beta) \text{ and } (y = x)\)

or \((x < \beta \land \exists \gamma \exists y : (x < \gamma < \beta) \land (g \in Y^\gamma) \land (g \text{ is a partial solution})\)

\(\land (y = g(x))\).

—in other words, the idea is to associate with \(x\) the value of a partial solution evaluated at \(x\). Note that when \(x < \beta\), there exists \(\gamma\) with \(x < \gamma < \beta\), since \(\beta\) is assumed to be a limit ordinal. For such a \(\gamma\) there exists, as we assumed that \([0, \beta] \subseteq A\) er antaget, a partial solution, and to this we associate the \(y\)-value \(g(x)\). By the uniqueness property of partial solutions, this \(y\)-value is given uniquely from \(x < \beta\).
Now use the axiom of substitution and determine a map \( g_0 \in Y^\beta \), such that \( y = g_0(x) \iff x < \beta \land P(x, y) \). For the transfinite sequence thus determined, \( g_0 \in Y^\beta \), it holds that for every \( x < \beta \) we can determine \( \gamma \) with \( x < \gamma < \beta \) and use this to realize that the restriction of \( g_0 \) to \([0, \gamma[\) is a partial solution. Therefore, \( g_0(x) = \Phi\left( (g_0(\xi))_{\xi < x} \right) \), i.e. \( g_0 \) is a partial solution. In great detail we have then shown that \( \beta \in A \) also in case \( \beta \) is a limit ordinal. By induction we can then conclude that \( A = [0, \alpha[ \). The last part of the proof is easy: Now we know that there exist partial solutions for every \( \beta < \alpha \). Then we can apply exactly the same kind of reasoning as above and conclude that there exists \( g \in Y^\alpha \) such that \( g(\beta) = \Phi\left( (g(x))_{x < \beta} \right) \) for all \( \beta < \alpha \). This is what we set out to prove.

The recursion theorem appears complicated at first sight. In fact, it only seldomly gives rise to problems. We use it to construct unique objects having certain prescribed properties. The essential condition for this to succeed is that at every step in the construction, i.e. for every ordinal number \( \beta \) we want to consider (in the theorem it was \(< \beta \)'s less than \( \alpha \)), it must be clear how to proceed based on the part of the construction already carried out. So, if only we know how to proceed, wherever in the process we are, it is certain that we will come to an end!

Often, when showing how to proceed, it is convenient to distinguish between the three cases corresponding to the three categories of ordinal numbers.

Now to a very useful variant of the recursion theorem which is characteristic by the fact that we do not from the start have to know the set \( Y \) used for the range of the transfinite sequences considered.

**Theorem 4.5.4** (Transfinite recursion over \( \alpha \), variant). Let \( \alpha \) be an ordinal number. Assume that to any \( \beta < \alpha \) and any transfinite sequence \( f \) of length \( \beta \) we have associated in a unique way some set \( \Phi(f) \). Then there exists a uniquely determined transfinite sequence \( f \) of length \( \alpha \), such that \( f(\beta) = \Phi\left( (f(x))_{x < \beta} \right) \) for every \( \beta < \alpha \).

The essential assumption means, in more detail, that there is given a predicate \( P(f, y) \) in two variables, such that the following three conditions are fulfilled:

1. \( \forall f \forall y \left( P(f, y) \Rightarrow \exists \beta \exists Y : \beta \in \mathbb{ON} \land \beta < \alpha \land f \in Y^\beta \right) \),
2. \( \forall f \left( \exists \beta \exists Y : \beta \in \mathbb{ON} \land \beta < \alpha \land f \in Y^\beta \Rightarrow \exists y : P(f, y) \right) \),
Axiomatic set theory

(3) \(\forall f \forall y_1 \forall y_2 \, P(f, y_1) \land P(f, y_2) \Rightarrow y_1 = y_2.\)

We then write \(y = \Phi(f)\) in place of \(P(f, y)\). Note that \(\Phi\) in the theorem is not a map (if it was, it would be defined on the set of all transfinite sequences of length less than \(\alpha\) – a collection of objects far too large to be a set. Anyhow, the object, \(f\) whose existence we assert, is a guinine set.

The postulated uniqueness in the theorem can be formulated more formally as follows:

\[ \exists \exists f \exists Y \left( f \in Y^\alpha \land \forall \beta < \alpha : P((f(x))_{x<\beta}, f(\beta)) \right). \]

For the formalist, we formulate the theorem as follows:

Let \(\alpha\) be an ordinal number and \(P\) a predicate in two variable. Assume that:

(i) \(\forall f \, (\exists y : P(f, y) \iff \exists \beta < \alpha \exists Y : f \in Y^\beta),\)
(ii) \(\forall f \forall y_1 \forall y_2 : P(f, y_1) \land P(f, y_2) \iff y_1 = y_2.\)

Then:

\[ \exists f \exists Y \left[ f \in Y^\alpha \land \forall \beta < \alpha : P(f_{|\beta}, f(\beta)) \right] \]

—and \(f\) is uniquely determined, i.e.:

\[ \exists f : f\text{ er en afbildning} \land \text{Dom}(f) = [0, \alpha[ \land \forall \beta < \alpha : P((f(x))_{x<\beta}, f(\beta)). \]

(Here, \(f_{|\beta}\) denotes the restriction of \(f\) to \(\beta = [0, \beta[\).

I think most will prefer the more loose formulation of the theorem. Regarding the proof, it follows very vlosely the proof of the first version of the recursion theorem. We only have to take a bit more care as we do not up front have a set \(Y\) to “steer after”. That we do not at any stage in the proof “run out of set theory” is ensured by applications of the axiom of substitution. Details are left to the interested reader.

We can go further and prove a third version of the recursion theorem, this time involving recursion over, not a set, but over all of \(\mathbb{ON}\!\!).

Theorem 4.5.5 (Transfinite recursion over \(\mathbb{ON}\!\!!\)). \(\dagger\) Assume that, to any ordinal number \(\alpha\) and any transfinite sequence \(f\) med \(\text{Dom}(f) = [0, \alpha[\) there is associated in a unique way a set \(\Phi(f)\). Then there exists a unique association \(\alpha \rightharpoonup f(\alpha)\), which, to every ordinal number \(\alpha\) associates a set \(f(\alpha)\), such that \(\forall \alpha \in \mathbb{ON} : f(\alpha) = \Phi\left( (f(x))_{x<\alpha} \right).\)

\(\dagger\) You can also allow induction over \(\mathbb{ON}\): If \(P(x)\) is a predicate, then from

\[ \forall \alpha \in \mathbb{ON} \left( \forall \beta < \alpha : P(\beta) \Rightarrow P(\alpha) \right) \]

you can conclude that \(\forall \alpha \in \mathbb{ON} : P(\alpha)\). (Easy!).
The reader is invited to give a more formal formulation of the result and to provide a proof.

4.6 Ordinal arithmetic

**Theorem 4.6.1** (Sum, product and exponentiation of ordinal numbers).

For every pair of ordinal numbers, \( \alpha \) and \( \beta \), there are uniquely defined ordinal numbers, denoted \( \alpha + \beta \), \( \alpha \cdot \beta \) and \( \alpha^\beta \) such that:

**Sum** \( \alpha + \beta \):
- \( \alpha + 0 = \alpha \) for \( \beta = 0 \),
- \( \alpha + S(\beta_0) = S(\alpha + \beta_0) \) for \( \beta = S(\beta_0) \),
- \( \alpha + \beta = \sup_{\gamma < \beta}(\alpha + \gamma) \) for \( \beta \) a limit ordinal.

**Product** \( \alpha \cdot \beta \):
- \( \alpha \cdot 0 = 0 \) for \( \beta = 0 \),
- \( \alpha \cdot S(\beta_0) = \alpha \cdot \beta_0 + \alpha \) for \( \beta = S(\beta_0) \),
- \( \alpha \cdot \beta = \sup_{\gamma < \beta}(\alpha \cdot \gamma) \) for \( \beta \) a limit ordinal.

**Exponentiation** \( \alpha^\beta \):
- \( \alpha^0 = 1 \) (\( = S(0) \)) for \( \beta = 0 \),
- \( \alpha^{S(\beta_0)} = \alpha^{\beta_0} \cdot \alpha \) for \( \beta = S(\beta_0) \),
- \( \alpha^\beta = \sup_{\gamma < \beta}(\alpha^\gamma) \) for \( \beta \) a limit ordinal.

The simple, though quite advanced proof is left to the reader! Just use recursion over \( \mathbb{ON} \) in a pretty obvious manner, starting with sum, then product and finally exponentiation. For the proofs start by keeping \( \alpha \) fixed and carry out the recursive construction over \( \beta \).

4.7 Natural numbers and the set of natural numbers

**Definition 4.7.1.** Let us agree to call \( x \) a **natural number** or 0, and let us write (for the time being purely symbolically) \( x \in N_0 \), if \( x \) is an ordinal number such that

\[
\forall y \ (y \leq x \Rightarrow y = 0 \lor \exists z : y = S(z)).
\]

Thus \( x \in N_0 \) means that \( x \in \mathbb{ON} \) and that \( x \) as well as every predecessor is a successor or 0.

Note that

\[
x \in N_0 \land x' \leq x \Rightarrow x' \in N_0, x \in N_0 \Rightarrow S(x) \in N_0.
\]

We write, again purely symbolically, \( x \in \mathbb{N} \) for \( x \in N_0 \land x \neq 0 \). If \( x \in \mathbb{N} \), \( x \) is a **natural number**.

**Definition 4.7.2.** A set \( x \) is **endelig**, if \( x = \emptyset \) or \( x \) is equipotent with a natural number.
More elegantly: $x$ is finite $\iff \exists n \in \mathbb{N}_0 : |x| = |n|$.

Finally we can get at the set of natural numbers:

**Theorem 4.7.3.** There is a set $\omega$, such that $n \in \omega \iff n$ is a natural number or $0$, i.e. $\omega = \{n|n$ is a natural number or $0\}$.

*Proof.* Superficially this looks easy: We just take an inductive set $x$—and such a set exists by the axiom of infinity—and then we use comprehension to define $\omega : \omega = \{n \in x|n$ is a natural number or $0\}$. Then $n \in \omega \Rightarrow n$ is a natural number or $0$. But the opposite implication is not obvious. The problem is to show that

$$n \text{ is a natural number or } 0 \Rightarrow n \in x. \tag{4.2}$$

This would be easy if we could apply induction over the natural numbers—but then we are in need of the set of natural numbers. Fact is that we have to use a rather awkward argument—which is only used this one time, as later we have access to genuine induction.

Thus, let $n$ be a natural number or $0$ and assume, with a view to an indirect argument, that $n \notin x$. Then $n \neq 0$. Therefore, there exists $n'$ such that $n = S(n')$. Then $n'$ cannot belong to $x$, hence $n' \in n\setminus x$ must hold. Thus, $n\setminus x \neq \emptyset$, and we may consider the first element, say $m$, in this set. Reasoning as above, we see that there exists $m'$, such that $m = S(m')$ and $m' \notin x$. But then $m'$ will be an element in $n\setminus x$, less than $m$, and this is impossible. We have reached a contradiction and conclude that $n \in x$. This ends the proof.

Now we can introduce $\mathbb{N}_0$ and $\mathbb{N}$ as notation for genuine sets—before they were just symbols—viz. $\mathbb{N}_0 = \omega$ and $\mathbb{N} = \mathbb{N}_0\setminus\{0\}$.

**Theorem 4.7.4.** $\omega$ is an ordinal number and can be characterized as the first limit ordinal.

*Proof.* Exploit (viii) in Theorem 4.4.2 and conclude that $\omega \in \mathbb{ON}$. Clearly, $\omega$ is not a natural number (or $0$), since then $\omega \in \omega$ would hold. We cannot either have that $\omega$ is a successor since, if $\omega = x \cup \{x\}$, $x \in \omega$ and then $\omega = S(x) \in \omega$ will hold. Therefore, $\omega$ is a limit ordinal. Clearly, it must be the smallest limit ordinal.

It is a key result that the set of natural numbers behaves “as it should”, i.e. the Peano axioms (Peano 1891) which are agreed to contain all essentials about the set of natural numbers hold true:

**Theorem 4.7.5** (Peanos axioms). The following holds:
4.7 Natural numbers and the set of natural numbers

(1) \(0 \in \omega \land \forall n \in \omega : 0 \neq S(n),\)
(2) \(\forall n \in \omega : S(n) \in \omega,\)
(3) \(\forall n \in \omega \forall m \in \omega : n \neq m \Rightarrow S(n) \neq S(m),\)
(4) \(\forall x \subseteq \omega \left[ (0 \in x \land \forall n \in x : S(n) \in x) \Rightarrow x = \omega \right].\)

**Proof.** (1), (2) and (3) are easy. For (4), assume that \(x \neq \omega\) and let \(\alpha\) be the first element in \(\omega \setminus x\). This element can neither be 0 or a successor (since, if \(\alpha = S(\beta), \beta \in x\) and then \(S(\beta) = \alpha \in x\) will hold). Then \(\alpha\) must be a limit ordinal < \(\omega\). This is impossible! We conclude that \(x = \omega\) must hold.

The result is important, perhaps especially in a historic light. Earlier one deduced all known structures and results about natural numbers based on Peano's axioms. This can be criticized as the basic fourth axiom incorporates quantization over subsets. As we have seen we may stay within first order logic provided we do not only look at the set of natural numbers but view this set as part of a much larger universe, indeed as a small corner of all of set theory.

It is easy, with the work done up to now, to introduce all basic structure on \(\omega\) (alias \(\mathbb{N}_0\)) and to establish all basic properties. In fact we have already done the essential parts as the ordering — indeed a well ordering — is in place and since ordinal arithmetic restricted to \(\omega\) gives us the basic algebraic structure of \(\omega\).

Our attitude is that with the set of natural numbers at hand all the other key objects of mathematics such as the sets \(\mathbb{Z}, \mathbb{Q}, \mathbb{R}\) and \(\mathbb{C}\) together with the essential structure associated with these sets can easily be constructed using clear and natural guiding principles. Though it involves quite a bit of work it is straightforward and does not involve new sophisticated ideas as was involved in the construction of the set of natural numbers. All this is supposed to be well known to the reader.
In this chapter we discuss the axiom of choice, AC. This axiom has a special place within ZFC in that it is not “definite” or “constructive” as the other axioms which assert the existence of certain sets. These other axioms lead – via the axiom of extentionality – to unique sets. Not so with AC.

5.1 AC – the Axiom of Choice

Let us state four variants of the axiom:

(i) Every indexed family of non-empty sets has a choice function:
If \((A_i)_{i \in I}\) is a family of non-empty sets, there exists a function, say \(f\), with \(I\) as domain such that \(f(i) \in A_i\) for every \(i \in I\),

(ii) every family of non-empty sets has a choice function:
if \(x\) is a set such that \(y \neq \emptyset\) for every \(y \in x\), then there exists a function, say \(f\), with \(x\) as domain such that \(f(y) \in y\) for every \(y \in x\),

(iii) any decomposition of a set has a choice function:
if \(y\) is a decomposition of a set \(x\) (i.e. \(\cup y = x\) and every set \(z \in y\) is non-empty and, furthermore, for \(z \in y\) and \(w \in y\) either \(z = w\) or \(z \cap w = \emptyset\)), there exists a function, say \(f\), with \(y\) as domain such that \(f(z) \in z\) for every \(z \in y\),

(iv) The power set of any set, with the empty set removed, has a choice function:
For any set \(x\) there exists a function, say \(f\), with \(\mathcal{P}(x)\setminus\{\emptyset\}\) as domain such that \(f(y) \in y\) for every \(y \in \mathcal{P}(x)\setminus\{\emptyset\}\).

**Theorem 5.1.1.** Within ZF (ZFC with AC removed) all four versions of AC given above are equivalent.
5.2 Some direct applications

Proof. Clearly, \((i) \Rightarrow (ii) \Rightarrow (iii)\). And \((iii) \Rightarrow (iv)\) follows by considering the decomposition of the disjoint union \(\bigsqcup \mathcal{P}\setminus\{\emptyset\}\) consisting of all sets of the form \(\{(y,z)\mid z \in y\}\) with \(y \in \mathcal{P}\setminus\{\emptyset\}\). Finally, \((iv) \Rightarrow (i)\) may be seen by considering to an indexed family of non-empty sets as appearing in \((i)\) the disjoint union \(A = \{(i,x)\mid i \in I, x \in A_i\}\). Then the desired choice function can be constructed from a choice function for \(\mathcal{P}\setminus\{\emptyset\}\). Details are left to the reader.

5.2 Some direct applications

Sometimes, AC can be applied directly. We give two instances of this:

**Theorem 5.2.1.** The countable union of countable sets is countable.

We leave the proof of this result to the reader, only noting that a well known argument makes the task easy provided each of the sets in the given countable family is represented as the range of a standard sequence. By AC you can reduce the problem to this situation.

And then the second direct application of AC which we want to point out:

**Theorem 5.2.2** (Vitali’s theorem). There is no countably additive, translation invariant measure defined on \(\mathcal{P}(\mathbb{R})\) and agreeing on the paving of intervals with the length function.

Before the proof, let us clarify the statement. It relates to desirable properties of a “length” function defined for subsets of \(\mathbb{R}\). Call the desired function \(\mu\). The requirements to \(\mu\) indicated in the theorem are those listed below, indicating by \(A\)‘s sets for which \(\mu\) are defined:

- \(\mu\) is defined on \(\mathcal{P}(\mathbb{R})\) and takes non-negative extended real values,
- \(\mu\) is finitely additive: \(\mu(\emptyset) = 0\) and, for any finite family \((A_k)_{k \leq n}\) of pairwise disjoint sets, \(\mu(\bigcup_{k \leq n} A_k) = \sum_{k \leq n} \mu(A_k)\),
- \(\mu\) is continuous: If \(A_k \uparrow A\) (\(A_1 \subseteq A_2 \subseteq \cdots\) and \(A = \bigcup_k A_k\)), then \(\mu(A_k) \rightarrow \mu(A)\),
- \(\mu\) is translation invariant: For any \(A\), for any \(t \in \mathbb{R}\), \(\mu(t + A) = \mu(A)\),
- For \(-\infty < a \leq b < \infty\), \(\mu([a,b]) = b - a\).

Most people with some mathematical background would think that such an object exists. The theorem states that in ZFC it does not.

Proof. Consider the decomposition of \(\mathbb{R}\) in equivalence classes under the equivalence relation \(\equiv \mod \mathbb{Q}\) \((t\ and\ s\ are\ equivalent\ if\ t\ can\ be\ expressed\ as\ an\ integer\ and\ \frac{t-s}{q}\ is\ an\ integer\ for\ some\ q\in\mathbb{N}\)
obtained from $s$ by adding a rational number). Intersect each equivalence class with the unit interval $[0, 1]$ and consider a choice function from the resulting collection of non-empty subsets of $\mathbb{R}$. Let $A$ be the range of this function.

Then, on the one hand $A$ must be “small” as there is an infinite collection of pairwise disjoint translates of $A$, viz. the collection of $t + A$ with $t \in \mathbb{Q} \cap [0, 1]$ which are all contained in a bounded set ($[0, 2]$ will do) and, on the other hand, $A$ must be “large” since some other such collection, viz. the collection of $t + A$ with $t \in \mathbb{Q}$ covers all of $\mathbb{R}$. Using the countable additivity of $\mu$ (the fact that for $(A_k)_{k \in \mathbb{N}}$ pairwise disjoint, $\mu(\bigcup A_k) = \sum \mu(A_k)$), the first fact shows that $\mu(A) = 0$, whereas the second fact shows that $\mu(A)$ must be positive. Contradiction!

This result – and related results of “non-existence” – I often refer to as the “necessity of measure theory”, as they make us embark on a large technical discipline meant to overcome the difficulties they point to. As one is not willing to give up AC, a closer analyses points to a solution where (i) above is given up and replaced by the weaker requirement that the set functions of interest only be defined on some subpaving, say $\mathcal{B}$ of $\mathcal{P}(X)$ ($X$ being the basic set you work in, above it was $\mathbb{R}$). Then $\mathcal{B}$ should be “structurally adequate” and “sufficiently rich”. The first requirement is taken care of by requiring that $\mathcal{B}$ be a Borel structure, or a $\sigma$-algebra as it is called by most authors, i.e. that $\mathcal{B}$ be $\bigcap$, $\bigcup$-closed. The second requirement is respected in various ways, depending on the nature of $X$ and the intentions of the investigations. A popular choice, which presupposes that there is given some (natural) topology on $X$, is to take for $\mathcal{B}$ the smallest $\sigma$-field containing the paving of open sets. This $\sigma$-field is called the Borel $\sigma$-field and is usually denoted $\mathcal{B}(X)$. One reason why we have spent some time on these issues is that the description of sets in a Borel structure such as $\mathcal{B}(X)$ is best understood by applying some of the more sophisticated parts of set theory. We shall return to this shortly. In fact, a good deal of what I will show in this section is connected in one way or another to problems of measure theory.

But first, we shall establish a general result, in fact a basic tool of set theory.
5.3 Zermelo’s well ordering principle and two immediate applications

Theorem 5.3.1 (Zermelo’s well ordering theorem). Any set can be well ordered.

Proof. Let $X$ be a set. We use a choice function $f$ on $\mathcal{P}(X) \setminus \{\emptyset\}$ to order the elements of $X$ in pretty much the same way as we naively attempted to do so, cf. Section 3.6. We define $(x_\alpha)_{\alpha \in \mathbb{ON}}$ by recursion over $\mathbb{ON}$ such that, for each ordinal number $\alpha$,

$$x_\alpha = \begin{cases} f(X \setminus \{x_\beta | \beta < \alpha\}) & \text{if } X \setminus \{x_\beta | \beta < \alpha\} \neq \emptyset \\ X & \text{if } X \setminus \{x_\beta | \beta < \alpha\} = \emptyset \end{cases}$$

Thus $x_\alpha \in S(X) = X \cup \{X\}$ for all $\alpha \in \mathbb{ON}$. In this construction, $X$ is a kind of “stop” symbol. Clearly, the process must “stop” at some stage, i.e. there must be some $\alpha$ such that $x_\alpha = X$. Otherwise we could use the substitution axiom for the property $P(x, \alpha)$ meaning that $x = x_\alpha$ and could conclude that $\mathbb{ON}$ is a set. Let $\alpha$ be the first ordinal number for which $x_\alpha = X$. Then clearly, $\beta \mapsto x_\beta$ is a bijection onto $[0, \alpha]$, hence $X$ can be well ordered.

As an important corollary to the proof we get:

Theorem 5.3.2. Any two sets, $X$ and $Y$ can be compared in the sense that either $|X| \leq |Y|$ or else $|Y| \leq |X|$.

Proof. Use the construction above, applied both to $X$ and to $Y$ and the fact that any two ordinal numbers can be compared.

It is only with results as the above that one gets full “benefit” from the introduction of ordinal- and cardinal numbers. We can (and did) define cardinal numbers without AC (as ordinal numbers which are first in their equipotence class) but without Zermelo’s theorem we cannot know that any set can be associated with a unique cardinal number. So only then can we think of $|X|$ as the cardinality of $X$, the least ordinal number equipotent with $X$. And only with AC is it possible to number all the infinite cardinal numbers by the ordinal numbers: $(\aleph_\alpha)_{\alpha \in \mathbb{ON}}$ – since how else could we know that there is a cardinal number larger than any given cardinal number? We would know through Cantor’s theorem that, given $X$, there is a set which is “larger” than $X$ (a set which $X$ can be embedded in but which is not equipotent to $X$).†

† the above discussion is a bit loose and possibly even misleading. Anyhow, a more
AC and some applications

Regarding the introduction of cardinal arithmetic and associated results, especially the result that for an infinite cardinal number $\kappa$, $\kappa \cdot \kappa = \kappa$, we refer to the exercises xxxx, considered an integral part of this section.

As an easy application of Zermelo’s theorem we mention the following very intuitive result which cannot, however, be proved in ZF alone:

**Theorem 5.3.3.** Every infinite set contains a countably infinite set.

The simple proof is left to the reader.

5.4 The Baire hierarchy

Now let us return a bit to the pavings of relevance to measure theory and show how they can be obtained in a constructive or semi-constructive way†. Let us work abstractly in a set $X$ provided with a “start paving” $\mathcal{B}_0$. For instance, we could have $X = \mathbb{R}^n$ and $\mathcal{B}_0 = G(\mathbb{R}^n)$, the paving of open subsets, or, for $n = 1$, $\mathcal{B}_0$ could be the paving of closed intervals.

Use transfinite recursion over $\text{On}$, to define, for each ordinal number $\alpha$, a paving $\mathcal{B}_\alpha$ in $X$ such that‡

- $\mathcal{B}_0$ is the given start-paving,
- if $\alpha$ is an odd ordinal number: $\alpha = \gamma + 2n + 1 = S\beta$ with $\beta = \gamma + 2n$ and $\gamma$ a limit ordinal or 0 (see footnote), then $\mathcal{B}_\alpha = \text{cl}_\cup \mathcal{B}_\beta$,
- if $\alpha$ is an even ordinal number: $\alpha = \gamma + 2n + 2 = S\beta$ with $\beta = \gamma + 2n + 1$ and $\gamma$ a limit ordinal or 0 (see footnote), then $\mathcal{B}_\alpha = \text{cl}_\cap \mathcal{B}_\beta$,
- if $\alpha$ is a limit ordinal, then $\mathcal{B}_\alpha = \bigcup_{\beta < \alpha} \mathcal{B}_\beta$.

From exercise xxxx we recall the special notation related to closures of pavings under certain set theoretical operations. We thus see that what we do is to start with the given start paving, then take the paving of all countable unions of sets in that paving, then the paving of all full discussion of what you can do without AC (related to Hartog’s number of a set) would have been appropriate – something to do for an eventual next version of LAM.

† again a full understanding and appreciation requires that we had developed a bit more of the theory of ordinal numbers without using AC (see previous footnote).
In particular, one will see that $\omega_1$, defined as the first uncountable ordinal number is a well defined ordinal number even without AC; i.e. in $ZF$.
‡ I should have defined before even and odd ordinal numbers. The even ones are of the form $\gamma + 2n + 2$, the odd ones of the form $\gamma + 2n + 1$ with $\gamma$ a limit ordinal or 0 and $n$ a natural number or 0. The decomposition of all successor ordinals in even and odd ones is easily established (but you need the axiom of foundation). Further results on the structure of ordinals, especially their so-called normal form is something to be considered for eventual next versions of LAM.
countable intersections of sets in the new paving obtained, then continue in that way alternating between the operations of countable unions and countable intersections until we get at a limit ordinal when we collect in one paving all sets met up to that point, continue with the alternating steps etc., etc.

The pavings \((\mathcal{B}_\alpha)_{\alpha \in \mathrm{ON}}\) constitute what is called the Baire hierarchy.

**Theorem 5.4.1.** (i) No new sets are added to the Baire hierarchy beyond \(\omega_1\), the first uncountable ordinal, i.e. \(\mathcal{B}_\alpha = \mathcal{B}_{\omega_1}\) for every \(\alpha > \omega_1\), (ii) \(\mathcal{B}_{\omega_1} = \text{cl}\cup\cap\mathcal{B}_0\), the smallest paving containing \(\mathcal{B}_0\) which is closed under countable unions and countable intersections, (iii) if \(\text{co}\mathcal{B}_0 \subseteq \text{cl}\cup\cap\mathcal{B}_0\), e.g. if every complement of a set in \(\mathcal{B}_0\) can be written as a countable intersection or as a countable union of sets in \(\mathcal{B}_0\), then \(\mathcal{B}_{\omega_1}\) is the \(\sigma\)-field generated by \(\mathcal{B}_0\), (iv) in “standard cases”, e.g. if \(X = \mathbb{R}\) and \(\mathcal{B}_0\) is the paving of closed intervals or if \(X = \mathbb{R}^n\) and \(\mathcal{B}_0\) is the paving of open sets (or the paving of closed sets), then new sets are added to the pavings considered for every ordinal number up to \(\omega_1\): \(\forall \alpha < \omega_1 : \mathcal{B}_\alpha \setminus \bigcup_{\beta < \alpha} \mathcal{B}_\beta \neq \emptyset\).

**Proof.** The first statement follows from the second. To prove (ii), consider a sequence \(A_1, A_2, \cdots\) of sets in \(\mathcal{B}_{\omega_1}\). Choose\(\dagger\) \(\alpha_1, \alpha_2, \cdots\), all less than \(\omega_1\) such that \(A_1 \in \mathcal{B}_{\alpha_1}, A_2 \in \mathcal{B}_{\alpha_2}, \cdots\). Put \(\alpha = \sup_{n \geq 1} \alpha_n\). Then \(\alpha < \omega_1\) (the key observation!). Clearly then, the union as well as the intersection of the sets \(A_n\) are both stones (i.e. members) in the paving \(\mathcal{B}_{\alpha+2}\). Thus \(\mathcal{B}_{\omega_1}\) is closed under the operations \(\cup\cap\) and \(\cap\cup\) and contains \(\mathcal{B}_0\). Clearly, it is the smallest such paving.

If \(\text{co}\mathcal{B}_0 \subseteq \text{cl}\cup\cap\mathcal{B}_0\), it is easy to prove by transfinite induction that \(\text{co}\mathcal{B}_\alpha \subseteq \text{cl}\cup\cap\mathcal{B}_0\) for every \(\alpha < \omega_1\). From this fact the third property follows.

We do not prove the last property here (though not too difficult – I indicated the idea in class).

\(\square\)

## 5.5 Zorn’s lemma

Many mathematicians do not feel all that comfortable with ordinal numbers and all that. I believe this will change, well is changing, in recognition of the importance for many parts of mathematics of arguments depending on properties of the ordinals. Anyhow, there are tools where

\(\dagger\) yes, there is an element of choice in the proof! But only the axiom of countable choice is required. For this reason we may consider the construction as providing only a “semi-constructive” characterization of the Borel sets.
the basic structure of the ordinals is hidden so that you can apply these tools without worrying about the finer details of set theory. One such result – and for the reasons mentioned a very popular result – is Zorn’s lemma:

**Theorem 5.5.1** (Zorn-s lemma). Every pre-ordered set for which every chain has an upper bound, has a maximal element.

**Proof.** Clarification first: We consider a pre-ordered set \( X = (X, \leq) \), i.e. \( \leq \) is reflexive and transitive. A chain in \( X \) is a subset \( X_0 \) which is totally ordered: \( \forall x, y \in X_0 : (x \leq y) \lor (y \leq x) \). So we assume that for every chain \( X_0 \) there is \( a \in X \) so that \( x \leq a \) for every \( x \in X_0 \). What we have to prove is that there is a maximal element in \( X \), i.e. an element \( x_0 \) such that, for every \( x \in X \), the implication \( x \geq x_0 \Rightarrow x \leq x_0 \) holds.

Then the proof: In fact, this is close to trivial, using Zermelo’s theorem: Given a well ordering of \( X \), say via a bijective map \( \alpha \curvearrowright x_\alpha \) of \([0, \alpha_0]\) onto \( X \), we just “travel” along the transfinite sequence \((x_\alpha)\), always making sure we travel along a chain and include all the elements in the chain we can. However, it is somewhat laborious to write this down formally. To do so, we define by transfinite recursion \((X_\alpha)_{\alpha \in \text{ON}}\) such that:

\[
X_\alpha = \begin{cases} 
\bigcup_{\beta < \alpha} X_\beta \cup \{x_\alpha\} & \text{if } \alpha < \alpha_0 \text{ and } x_\alpha \text{ is an upper bound of } \bigcup_{\beta < \alpha} X_\beta, \\
\bigcup_{\beta < \alpha} X_\beta & \text{otherwise .}
\end{cases}
\]

Clearly, the sets \( X_\alpha \) are eventually constant, in fact eventually equal to \( X_{\alpha_0} \). This set, denote it by \( X^* \), is a chain, hence has an upper bound. Consider any upper bound \( x^* \) of \( X^* \). We claim that \( x^* \) is a maximal element. To see this, let \( x \in X \) and assume that \( x \geq x^* \). We have to prove that \( x \leq x^* \). First determine the index \( \alpha \) of \( x \): \( x_\alpha = x \). Since \( x \geq x^* \) and \( x^* \) is an upper bound for \( X^* \), hence also for \( \bigcup_{\beta < \alpha} X_\beta \), \( x \) is an upper bound of \( \bigcup_{\beta < \alpha} X_\beta \), hence is included in \( X_\alpha \). As \( x^* \) is an upper bound for this set, \( x^* \geq x \), as desired.

Hopefully the reader will agree that the proof is close to trivial – and very easy to explain on the blackboard. The fact that quite some formal writing was necessary to document this is indicative of other instances where an application of Zorn’s lemma works very elegantly but often, so is my opinion, an application of Zermelo’s theorem tells more about

† if we add the property of anti-symmetry \(((x \leq y) \land (y \leq x) \Rightarrow x = y)\) we get a partially ordered set. However, watch out for slightly different usage in the literature.
5.6 Some applications of Zorn’s lemma

Let $X$ be a vector space over the field $F$. Recall that a set $E \subseteq X$, often written in indexed form $(e_i)_{i \in I}$, is a basis for $X$ if every $x \in X$ has a unique representation $x = \sum_{e \in E} \lambda_e e$ (alternatively $x = \sum_{i \in I} \lambda_i e_i$) with scalars -- the $\lambda$’s -- from $F$ and $\lambda_e = 0$ except for finitely many $e \in E$ (alternatively, $\lambda_i = 0$ except for finitely many indices $i$).

Theorem 5.6.1. Every vector space has a basis.

Proof. Consider a maximal linearly independent subset. Clearly, this must be a basis. □

In particular, $\mathbb{R}$ viewed as a vector space over the rationals, has a basis, a so-called Hamel basis. Let $E \subseteq \mathbb{R}$ be such a basis. We may use this to construct “strange” solutions of Cauchy’s functional equation:

$$f(x + y) = f(x) + f(y).$$

Here, $f$ is a real function of a real variable and we search for such functions satisfying the above equation for all real $x,y$. Of course, a homothetic transformation, $x \mapsto ax$ with $a$ a fixed real constant is a solution. The question is if there are others. Indeed there is: Choose any basis vector $e_0 \in E$ and consider $f$ given by

$$f(x) = \lambda e_0(x)$$

with $x = \sum_{e \in E} \lambda_e e$. If this function was of the form $x \mapsto ax$, on the one hand, $a \neq 0$ (consider $f(e_0)$) and, on the other, $a = 0$ must hold (consider $f$ applied to some other basis vector).

Let us turn to a general result of great applicability, the Hahn-Banach theorem. We formulate the basic variant of results of this type. The setting is a real vector space $X$, i.e. the field of scalars is the reals. What we investigate is linear functionals which are linear mappings into the simplest such vector space, the reals itself. Though we shall not pursue this, we are now again touching results related to measure theory. Fact is that such functionals are often provided by a measure through a map of the type $x \mapsto \int x \, d\mu$. Typically then, the vector space will be a function space -- so the $x$’s will be functions defined on some underlying set -- and
\[ AC \text{ and some applications} \]

\[ \mu \text{ will be a measure on some suitable Borel structure on the underlying set.} \]

For our purposes, let us say that a map \( x \mapsto \|x\| \) defined on the real vector space \( X \) and having non-negative values is of \textit{norm type} if \( \|x + y\| \leq \|x\| + \|y\| \) for \( x, y \in X \) and if \( \|ax\| = a\|x\| \) for \( x \in X, \ a \geq 0 \).

**Theorem 5.6.2** (Hahn-Banach theorem). Let \( X \) be a real vector space, \( x \mapsto \|x\| \) a map of norm-type, \( X_0 \) a subspace of \( X \) and \( f_0 : X_0 \to \mathbb{R} \) a linear functional defined on \( X_0 \) and on that subspace dominated by \( \| \cdot \| \) \( (f(x) \leq \|x\| \text{ for } x \in X_0) \). Then there exists an extension of \( f_0 \) to a linear functional \( f \) defined on all of \( X \) and dominated there by \( \| \cdot \| \).

**Proof.** Let \( f \) be a maximal extension of \( f_0 \) to a linear functional dominated on the subspace of definition, call it \( X' \), by \( \| \cdot \| \). The precise meaning of this statement is pretty obvious and the fact that such an extension exists is proved by a straightforward application of Zorn’s lemma. All that remains to prove is that \( X' = X \). Suppose not, and let \( u \in X \setminus X' \). For any \( c \in \mathbb{R} \), let \( f_c \) denote that extension of \( f \) to a linear functional on the vector space generated by \( X' \) and \( u \), which is defined by taking \( f_c(u) = c \). We shall arrive at a contradiction by showing that \( c \) may be chosen so that \( f_c \) also satisfies the condition of domination. The requirement for this is that, for all \( a > 0 \) it holds that for all \( x \) and \( y \) in \( X' \) or, more conveniently, for all \( ax \) and \( ay \) in \( X' \), that the two inequalities \( f_c(ax - au) \leq \|ax - au\| \) and \( f_c(ay + au) \leq \|ay + au\| \) hold. These inequalities are independent of \( a > 0 \) and amount to the inequalities

\[
\begin{align*}
    f_0(x) - c &\leq \|x - u\| \text{ and } f_0(y) + c \leq \|y + u\|.
\end{align*}
\]

Thus, an appropriate value for \( c \) can be chosen if and only if

\[
\sup_{x \in X'} \left( f_0(x) - \|x - u\| \right) \leq \inf_{y \in X'} \left( \|y + u\| - f_0(y) \right).
\]

Clearly, this inequality holds as, for \( x, y \in X' \),

\[
f_0(x) + f_0(y) = f_0(x+y) \leq \|x+y\| = \|(x-u)+(y+u)\| \leq \|x-u\| + \|y+u\|.
\]

All things considered, we have arrived at the desired contradiction and the theorem is proved. \( \square \)

We note that for this proof, the “deep part” (the part where AC was used) was a matter of one or two lines to prove, whereas the “elementary” part took up some more space (and if we had gone to the case of vector spaces over \( \mathbb{C} \) – for which an appropriate version of the result also holds –
we would have needed substantially more reasoning for the “elementary” part).

Another feature is also worth while stressing: Often, as we have seen, when arguments involving AC are involved, we are led to “weird” objects. Not so in this case. Fact is that the important property of domination protects us again “weird” objects as domination normally implies continuity w.r.t. topologies one is interested in.

5.7 Equivalences in ZF

Having had good use by now of either AC directly or some of the consequences of AC it lies nearby to examine the strength of these consequences as measured against the original axiom, AC. The fact is that there is no difference “modulo ZF”. What we mean is that the following result holds:

**Theorem 5.7.1.** Assume that you work in a model of set theory for which the axioms of ZF hold. In such a model the following statements are equivalent:

(i) AC

(ii) Zermelo’s well ordering theorem

(iii) Zorn’s lemma.

Proof. We have seen that (i) \(\Rightarrow\) (ii) \(\Rightarrow\) (iii) and from (iii) to (i) is easy: Given some family of non-empty sets, consider partial choice functions defined in the natural way and order them by extension. Use Zorn’s lemma to find a maximal choice function and note that any such maximal choice function must be defined on the domain of all non-empty sets in the family.

Other interesting properties have the same strength – modulo ZF – as the properties singled out. This holds, in particular of the comparison property of Theorem xxxx. The proof of this fact is not quite trivial and will not be given here.

Other properties, such as the Hahn-Banach theorem have a lot of the flavour of AC but are not quite as strong. We shall not enter into a detailed discussion of these type of results.

5.8 Inadequacy of sequences for general topology

First some comments, not all that easy to understand if you have no prior experience with general topological spaces.
The aim of general topology is to provide a framework for the study of continuity and convergence. It is generally agreed today that an appropriate starting point is the not-so-intuitive notion of a topological space which is nothing but a paved set \((X, \mathcal{G})\) with \(\mathcal{G} = \mathcal{G}(X)\) a paving containing \(\emptyset\) and \(X\) and being closed under the operations \(\cup, \cap\) and \(f\). The sets (stones) in the paving \(\mathcal{G}\) are called open sets and typically denoted by the letter \(G\). The paving \(F = \text{co}\mathcal{G}\) is the paving of closed sets, and these sets are typically denoted by the letter \(F\). A neighbourhood of a point \(x \in X\) is a set containing an open set which contains \(x\). Typical notation: \(N(x)\).

Convergence \(x_n \to x\) for a sequence in a topological space is defined to mean that

\[
\forall N(x) : x_n \in N(x) \text{ ev}.
\]

Here, “ev” is short for “eventually”. Thus, in more detail, the statements is:

\[
\forall N(x) \exists n_0 \forall n \geq n_0 : x_n \in N(x).
\]

A simple and intuitively very natural definition.

Continuity of a map \(f : X \to Y\) between topological spaces can be defined to mean, as suggestions of appropriate definitions, either of the following properties:

(i) \(f^{-1}(G(Y)) \subseteq G(X)\),

(ii) \(\forall x \forall N(f(x)) \exists N(x) : f(N(x)) \subseteq N(f(x))\),

(iii) \(\forall x_n \to x : f(x_n) \to f(x)\).

re (i): An elegant but highly unintuitive definition!

re (ii): Just the classical \(\varepsilon, \delta\) definition. Much more intuitive.

re (iii): The most intuitive definition: Continuity as the stability property: preservation of convergent sequences.

Conditions (i) and (ii) are equivalent and they imply (iii). But in general, (iii) does not imply the other conditions. Furthermore, when it does, e.g., in all metrizable spaces, in particular in all Euclidean spaces (a fact all readers of these notes should really be familiar with), we observe, thinking more over the natural (indirect) proof, that a choice is involved. This element of choice cannot be avoided. What one needs is the axiom of countable choice.

I am convinced that many students will benefit a lot from more focus on sequences and their generalizations than is common practice in almost every textbook on general topology. Sequential concepts are mostly more intuitive and the proofs suggest themselves more easily. However,
a certain degree of maturity as acquired through a course like this one is necessary.

Let us, in a loose discussion, focus on two properties: The general and natural property of a closed set as one from which you cannot escape by convergence: $F$ is closed ought to be equivalent with the implication $(\forall n : x_n \in F) \land (x_n \to x) \Rightarrow x \in F$. This is OK in Euclidean spaces but not in general. As an indication of an example consider the topology of pointwise convergence for real functions of a real variable and consider the set of indicator functions of countable subsets of $\mathbb{R}$. Clearly, you cannot escape from this set by taking limits of convergent sequences. But any indicator function ought to be included in the set for it to deserve the property of being closed (I indicated in class why this is really so). So sequences do not suffice.

Another property is the Bolzano-Weierstrass property and similar statements: The important possibility to extract convergent subsequences from bounded sequences (in Euclidean spaces). This means that from a sequence we can so to speak choose for a given set if we want to be close to the set or close to the subset. The property is closely related to the following result, I think first pointed out by Cantor:

**Theorem 5.8.1.** Let $(x_n)$ be a sequence in $X$. Then, for every countable paving $\mathcal{E}$ on $X$, there exists a subsequence $(x_{n(k)})$ such that, for each $E$ in $\mathcal{E}$, either $x_{n(k)}$ lies eventually in $E$ or else, $x_{n(k)}$ lies eventually in the complement of $E$.

For a proof, see TOP, Theorem 3.4.

In order to discuss general compactness results, results with similar features as the Bolzano-Weierstrass theorem it turns out to be important to generalize the above result by allowing that the chosen subsequence “chooses side” for more sets than just sets in a countable paving. But really the above result cannot be improved in that direction.

The solution to the problems indicated lies in allowing a generalized notion of sequence, called either generalized sequence or net.

5.9 Generalized sequences, nets

The notion we need singles out the key properties needed of the index set in a sequence. It turns out that we may replace $\mathbb{N}$ by far more general types of sets: By a directed set $D$ we understand a preordered set such that every finite set has an upper bound.

Clearly the usual index set used for sequences, $\mathbb{N}$, is a directed set
in the usual ordering. Also, any set $D$ in the diffuse ordering ($\alpha \leq \beta$ for any $\alpha, \beta \in D$) is a directed set (and a very special one, pointing in “all kind of directions”). In class I indicated examples related to the definition of Riemann sums for the definition of classical integrals (and also for the definition of not-so-classical integrals associated with the names of Kurzweil and Henstock). Here, let me mention one more and a quite important example:

**Example 1** Let $X$ be a set and $\mathcal{U}$ an $\cap$-closed paving on $X$ consisting of non-empty sets. Consider the set $D = \{(U, x) | x \in U \wedge U \in \mathcal{U}\}$ and order (preorder!) this set by agreeing that for $\alpha = (U, x) \in D$ and $\beta = (V, y) \in D$, $\beta \geq \alpha$ shall mean that $V \subseteq U$. Clearly the preorder is reflexive and transitive. If $(U_k, x_k) \in D$ for $k \leq n$, we have that $(\bigcap_{k \leq n} U_k) \neq \emptyset$, hence we can choose $x$ so that $(\bigcap_{k \leq n} U_k, x) \in D$ and this element is an upper bound for the finitely many elements we started with. Thus $D$ is indeed a directed set.

Let $X$ be a set. A net in $X$, or a generalized sequence in $X$, is a map defined on a directed set and having values in $X$. Let $\varphi : D \to X$ be a net. Very often we write the net as $(x_\alpha)_{\alpha \in D}$ or just as $(x_\alpha)$. Then $x_\alpha = \varphi(\alpha)$ for $\alpha$ in $D$ and $D$ is the index set of the net.

The examples considered or indicated above give rise to certain nets. When the index set is $\mathbb{N}$, we regain the usual notion of a sequence. By the diffuse net in $X$ we understand the identity map on $X$ with the domain directed by the diffuse ordering. The nets related to Riemann sums I shall not comment on. But there is an important construction related to the directed set of Example 1, viz. the net defined on $D$ by the projection map $(U, x) \mapsto x$. This we call the natural net associated with the paving $\mathcal{U}$.

Now, in relation to Theorem xxx, we aim at a surprising and far reaching variant which depends critically on the extension given of the notion of a generalized sequence, a net, and start with a definition: The net $(x_\alpha)$ is a universal net in $X$ if, for every subset $A$ of $X$, the net is in $A$ eventually, or in $\overline{A}$ eventually. Equivalently, we may define a universal net by the requirement that, for every subset $A$ of $X$, $x_\alpha$ is in $A$ eventually if it is in $A$ frequently ($x_\alpha \in A$, frequently means that $\forall \alpha \in D \exists \beta \geq \alpha : x_\beta \in A$). The only example of a universal net which is known in concrete terms is that of a net which is eventually constant.

Next we introduce the important notion of a subnet. Let $\varphi : D \to X$ be a net in $X$ and let $I$ be a directed set. Then $\sigma : I \to X$ is a subnet of
ϕ if there exists a mapping \( \psi : I \rightarrow D \) such that \( \sigma = \varphi \circ \psi \) and such that, for every \( \beta \) in \( D \), \( \psi(i) \geq \beta \) eventually \([i]\). In our usual terminology with \( x_\alpha = \varphi(\alpha) \), the definition can be expressed as follows: The net \((x_{\psi(i)})_{i \in I}\) is a subnet of \((x_\alpha)_{\alpha \in A}\) if, for every \( \beta \) in \( D \), \( \psi(i) \geq \beta \) eventually \([i]\).

**Theorem 5.9.1.** Every net has a universal subnet.

**Proof.** Consider the paving \( F \) of all sets which contains a tail set, a set of the form \( \{x_\alpha | \alpha \geq \alpha_0\} \) for some \( \alpha_0 \in D \). This paving is a filter in \( X \) which, by definition, is a paving in \( X \) closed under finite intersections which does not contain the empty set and contains every set which contains a set in the paving. Use Zorn’s lemma and construct a maximal filter, call it \( U \), which contains \( F \). Then, for every \( A \subseteq X \), either \( A \in U \) or \( \bar{A} \in U \). To see this, assume that \( A \notin U \) and consider the paving of all sets \( B \) with \( A \cup B \in U \). By maximality of \( U \), this paving coincides with \( U \), thus, in particular, \( \bar{A} \in U \). Now consider the directed set \( I \) of all pairs \((U, \alpha)\) with \( x_\alpha \in U \in U \). By consideration of the composed map: \((U, \alpha) \leadsto \alpha \leadsto x_\alpha \) we obtain the desired universal subnet of the given net. Details to check are left to the reader.

Admittedly, the full understanding and appreciation of the results developed requires a bit of acquaintance with the new ideas, especially the development of further parts of general topology (as indicated, preferably in a way which is not exactly the way the subject is taught at most universities, alas, including ours).

A bit simpler than the developments above related to universal nets is the introduction of the natural extended notion of convergence. Indeed, if \((X, \mathcal{G})\) is a topological space, \( x \) a point in \( X \) and \((x_\alpha)_{\alpha \in D} \) a net in \( X \) then we say that the net **converges to** \( x \) and write \( x_\alpha \rightarrow x \) if \( \forall N(x) : x_\alpha \in N(x), ev. \) With this extension it holds quite generally that a map between topological spaces is continuous if and only if it preserves convergent nets. Precise formulation and proof, this time not relying on any version of AC, is left to the reader.