Set theory

Giulio Lo Monaco

Academic year 2019/2020, second semester
Masaryk University, Brno

Contents

1   Basics of first-order logic

The first lecture will be dedicated to the fundamentals of first-order logic, so that we have the means to understand what it means to formulate a mathematical statement in a rigorous manner, and what it means that such a statement is satisfied. With this in mind, we will proceed to all the structures, generally called sets, that satisfy all the axioms of ZFC axiomatic system of set theory.

1.1   Formal languages

A language is the collection of the following data:

- a non-empty set of sorts, denoted as s_i;

- symbols for variables, each assigned to a sort;

- the symbols for equality $\approx$, disjunction $\vee$, conjunction $\wedge$, negation $\neg$, implication $\rightarrow$;

- the existential quantifier $\exists$ and the universal quantifier $\forall$;

- for each $n \geq 0$, symbols for n-ary predicates, equipped with an assignment of sorts $s_1, \ldots, s_n$;

- for each $n \geq 0$, symbols for n-ary functions, equipped with an assignment of sorts $s_1, \ldots, s_n$ and an additional sort s.

Remark 1.1. When we say "set", for the moment, we don't mean the same thing as the object of study of set theory. These are rather just naive collections of things, with no axioms to be respected and no real mathematical content. The context in which we use the word "set" with this meaning is in fact often called naive set theory, as opposed to axiomatic set theory.

Remark 1.2. In writing s_1, s_2 etc., one shouldn't think of subscript numbers as actual numbers, rather as pure symbols used with no other purpose than giving names to sorts (or variables, or functions...)

Definition 1.3. *For $n = 0$, an n-ary (or nullary) function symbol is also called a constant symbol.*

1.2 Terms

Our next step will be to inductively define terms in a given language. Terms are particular strings of symbols that can be thought as representing the objects about which formal statements can be formulated. Every term comes with an assignment of a sort.

Definition 1.4. *Fixed a language $\mathcal{L}$, the collection of terms is inductively defined by:*

1. *every constant or variable symbol is a term of the corresponding sort;*

2. *if $t_1, \ldots, t_n$ are terms of respective sorts $s_1, \ldots, s_n$ and f is a function of sorts $s_1, \ldots, s_n, s$, then the string $ft_1 \ldots t_n$ is a term of sort s.*

1.3 Well-formed formulas

Finally, well-formed formulas, or wff's for short, are those string of symbols in the given language that are meant to represent what we commonly call mathematical statements. They can also be defined inductively, starting from the already known terms, and using the remaining part of the language, which we didn't use in order to define terms.

Definition 1.5. *Atomic formulas are strings of one of the following forms:*

1. *$t_1 \approx t_2$, where t_1 and t_2 are terms;*

2. *$Pt_1 \ldots t_n$ where P is a symbol for predicate of sorts $s_1, \ldots, s_n$ and $t_1, \ldots t_n$ are terms such that t_i is of sort s_i.*

Definition 1.6. *Well-formed formulas (of first-order, finitary logic) are defined inductively by the following steps:*

1. *all atomic formulas are wff's;*

2. *if ϕ and ψ are wff's, then so are $\neg\phi$ and $\phi \to \psi$;*

3. *if Θ is a finite set of wff's, then $\vee\Theta$ and $\wedge\Theta$ are wff's;*

4. *if ϕ is a wff and x is a symbol for variable, then $\exists x\phi$ and $\forall x\phi$ are wff's.*

1.4 Interpretations

So far, terms and well-formed formulas are just meaningless strings of symbols, satisfying the only condition that they have been obtained as explained in the previous sections. We would like to associate some meaning to such strings, and subsequently to be able to decide whether the statement that they represent is true or false.
First, we will say what an interpretation is, without attaching any judgment of truthfulness or falsehood to its outcomes for the moment.

Let us fix a language $\mathcal{L}$. We will describe an interpretation of its basic components, and then we will move on to interpret terms and wff's as well.

- To every sort s, we associate a set $I(s)$ (again, in the naive sense).

- To every n-ary predicate symbol P of sorts $s_1, \ldots, s_n$ we associate a subset $I(P) \subseteq I(s_1) \times \ldots \times I(s_n)$.

- Finally, to every n-ary function symbol f of sorts $s_1, \ldots, s_n, s$ we associate a function $I(f) : I(s_1) \times \ldots \times I(s_n) \to I(s)$.

Now we need some notation. For a sequence $\bar{x} = (x_1, \ldots, x_n)$ of variables, where x_i is of sort s_i, we write $\mathsf{C}_{\bar{x}}$ for $I(s_1) \times \ldots \times I(s_n)$, and we call it the "context of variables $\bar{x}$". Now the idea is that a term t of sort s will be interpreted as a function $\mathsf{C}_{\bar{x}} \to I(s)$, where $\bar{x}$ is the sequence of all variable symbols appearing at least once in the explicit formulation of t. Heuristically, we should think of this as taking every possible choice of values for the given variables to the result of plugging this choice in the definition of t. More explicitly, this is done as follows:

1. if $t = x$ of sort s, then $I(t) = \mathrm{id} : I(s) \to I(s)$;

2. if $t = ft_1 \ldots t_n$ with t_i of sort s_i, then $I(t)$ is the composite

$$\mathsf{C}_{\bar{x}} \xrightarrow{(I(t_i))_{i=1}^n} \prod_{i=1}^n I(s_i) \xrightarrow{I(f)} I(s).$$

A formula ϕ will be interpreted as a subset of $\mathsf{C}_{\bar{x}}$, where $\mathsf{C}_{\bar{x}}$ is now the context of all variables appearing in the formulation of ϕ. Heuristically, it selects all the possible choices for the given variables that make ϕ hold once substituted to the respective variable symbols in it.

- Given terms t_1 and t_2, then $I(t_1 \approx t_2) = \{\bar{a} \in \mathsf{C}_{\bar{x}} | I(t_1)(\bar{a}) = I(t_2)(\bar{a})\}$;

- For $t = Pt_1 \ldots t_n$, with t_i of sort s_i, then $I(t) = \{\bar{a} \in \mathsf{C}_{\bar{x}} | g(\bar{a}) \in I(P)\}$, where $g = (I(t_i))_{i=1}^n : \mathsf{C}_{\bar{x}} \to \prod_{i=1}^n I(s_i)$;

- $I(\neg\phi) = \mathsf{C}_{\bar{x}} \setminus I(\phi)$;

- $I(\vee_{j \in J}\phi_j) = \cup_{j \in J}I(\phi_j)$;

- $I(\wedge_{j \in J}\phi_j) = \cap_{j \in J}I(\phi_j)$;

- $I(\phi \to \psi) = (\mathsf{C}_{\bar{x}} \setminus I(\phi)) \cup I(\psi)$

Remark 1.7. In order to understand the interpretation of the implication symbol, remember that the expression $\phi \to \psi$ is logically equivalent to $\neg\phi \vee \psi$

Finally, we interpret the quantifiers as follows. Observe that, if y is a variable of sort s', not necessarily among the variables of $\bar{x}$, if variable y appears in ϕ, then $I(\phi) \subseteq \mathsf{C}_{\bar{x},y}$, whereas we want to find an interpretation of $\exists y\phi$ and $\forall y\phi$ as subsets of $\mathsf{C}_{\bar{x}}$.

- $I(\exists y\phi) = \{\bar{a} \in \mathsf{C}_{\bar{x}}|$ there exists $b \in I(s')$ such that $(\bar{a}, b) \in I(\phi)\}$;

- $I(\forall y\phi) = \{\bar{a} \in \mathsf{C}_{\bar{x}}|$ for all $b \in I(s')$ we have that $(\bar{a}, b) \in I(\phi)\}$.

1.5   Truth

Now we turn our attention to deciding whether a formula is true under a given interpretation. We use the machinery of sequent calculus, which apparently brings in an additional complication, in that it introduces a new symbol to the pile that we already have, but in fact offers a slick and elegant way of defining truth.

Definition 1.8. *A sequent is a string of the form $\Theta \Rightarrow \phi$, where Θ is a finite set of formulas, ϕ is a single formula and $\Rightarrow$ is a new symbol, not comprised in the language.*

We will say that a sequent of the form above is true under a given interpretation I if, being $\mathsf{C}_{\bar{x}}$ the context of all variables in Θ and ϕ, then $I(\wedge\Theta) \subseteq I(\phi)$ as subsets of $\mathsf{C}_{\bar{x}}$.
An closer inspection will reveal that this definition of truth reflects in fact the intuitive notion of it that we commonly have (think: if all the premises in the antecedent Θ are true, then we conclude that ϕ is also true).

2   Some consequences of the axioms of ZFC

The purpose is to understand what a theory is and what it means to model it. We will then narrow our attention to the case of ZFC set theory.

2.1   ZFC set theory, a starter

Definition 2.1.   • *A theory is a pair $(\mathcal{L}, A)$, where $\mathcal{L}$ is a formal language and A is a set of sequents, called the axioms of the theory.*

- *A model for a theory is an interpretation I making all its axioms true.*

Remark 2.2. Up to this point, whenever we used the word "set" we were doing so in a naive sense. The sets which are used in order to define an interpretation, or the set of axioms of a theory, are all objects of naive set theory.

Sets in a more rigorous sense are elements of a model for an axiomatic set theory, e.g. ZFC.

Definition 2.3. *Let x be a variable symbol and ϕ a formula in which x appears. Then x is said to be free in ϕ if its first occurrence is not quantified, i.e. it is not preceded by $\forall$ or $\exists$.*

Given a formula $\phi(x_1, \ldots, x_n)$ where the variables $x_i's$ are free, and given other $a_1, \ldots, a_n$ variable symbols, then $\phi(a_1, \ldots, a_n)$ will denote the formula ϕ where all the occurrences of the variable symbols $x_i's$ have been replaced by the corresponding symbols $a_i's$.

Example 2.4. ZFC set theory has a language containing a binary relation symbol $\in$ and a constant symbol $\emptyset$. We use the following shortcuts:

- $\exists! x \phi(x) \leftrightarrow [\exists x \phi(x) \wedge \forall(\phi(y) \rightarrow y \approx x)$

- $x \approx y \cup z \leftrightarrow \forall w[w \in x \leftrightarrow (w \in y \vee w \in z)]$

- $x \subseteq y \leftrightarrow \forall z(z \in x \rightarrow z \in y)$

Omitting the axioms of regularity and choice for the moment, the remaining part of the theory is:

- **Axiom of extensionality** $\forall x \forall y \forall z(z \in x \leftrightarrow z \in y) \leftrightarrow x \approx y$

- **Axiom schema of separation** Given a well-formed formula $\phi(x)$ where x is free, then there is an axiom $\forall z \exists y a \in y \leftrightarrow (a \in z \wedge \phi(a))$

- **Axiom of pairing** $\forall x \forall y \exists z(x \in z \wedge y \in z)$

- **Axiom of union** $\forall F \exists A \forall x \forall y(x \in y \wedge y \in F) \rightarrow x \in A$

- **Axiom schema of replacement** Given a well-formed formula $\phi(x, y, A)$ where x, y and A are free, then there is an axiom
  $\forall A([\forall x \in A \exists! y \phi(x, y, A)] \rightarrow \exists B \forall y[y \in B \leftrightarrow \exists x \phi(x, y, A)])$

- **Axiom of infinity** $\exists x[\emptyset \in x \wedge \forall y(y \in x \rightarrow y \cup \{y\} \in x)]$

- **Axiom of power set** $\forall x \exists y \forall z(z \subseteq y \rightarrow z \in y)$

Remark 2.5. Separation and replacement are not axioms, rather families of axioms, indexed by the (naive) set of formulas in our language. ZFC set theory is thereforse comprised of an infinite set of axioms.

Example 2.6. Given a family of sets $(z_i)_{i \in I}$, then $\wedge_{i \in I}(x \in z_i)$ is a well-formed formula, therefore we can define intersection by choosing one z_j and then setting $\cap_{i \in I} z_i = \{x \in z_j | \wedge_{i \in I} x \in z_i\}$, where we used an axiom of separation.

Note that, although $\vee_{i \in I}(x \in z_i)$ is also a well-formed formula, we cannot similarly define union by using separation, because there is a priori no set containing all of $z_i's$ where we can apply our formula. This is precisely why we need the axiom of union.

2.2 A couple of important theorems

Remember that we say that two sets have the same cardinality, and we write $|A| = |B|$, if there is a bijective function between them. We also write $|A| \leq |B|$ if there is an injective function $A \to B$, and $|A| < |B|$ will then be an abbreviation for $|A| \leq |B| \wedge |A| \neq |B|$. We denote the power set of A, whose existence is ensured by the corresponding axiom, as $\mathcal{P}(A)$.

Theorem 2.7 (Cantor's theorem). *Given a set A, then $|A| < |\mathcal{P}(A)|$.*

Proof. The function $A \to \mathcal{P}(A)$ defined by $a \mapsto \{a\}$ is evidently injective, therefore $|A| \leq |\mathcal{P}(A)|$. We now show that no map from A to its power set can be surjective.
Consider an arbitrary map $f : A \to \mathcal{P}(A)$, and use separation to define a subset $X = \{a \in A | a \notin f(a)\}$. Suppose there is $b \in A$ such that $f(b) = X$. Then by examining the definition we will obtain that $b \in X \Leftrightarrow b \notin X$, which is a contradiction. It is then impossible to have such a b, so f is not surjective. $\square$

Theorem 2.8 (Cantor-Bernstein theorem). *Given two sets A and B, suppose there are two injective functions $A \to B$ and $B \to A$. Then $|A| = |B|$.*

This is proven in the theory section of the course. In particular, the proof uses the following

Theorem 2.9 (Tarski's fixed point theorem). *Let X be a complete lattice and $f : X \to X$ an order-preserving function. Then there is a fixed point for f, i.e. $\exists x f(x) = x$.*

Proof. By separation, define $Y = \{y \in X | y \leq f(y)\}$. Now take $x = sup Y$ which exists because X is complete. We have then that $\forall y \in Y, y \leq f(y) \leq f(x)$ because f is order-preserving, therefore by definition of supremum we obtain $x \leq f(x)$.
Moreover, this relation implies, again since f is order-preserving, that $f(x) \leq f^2(x)$, which means that $f(x) \in Y$. Consequently, $f(x) \leq x$. Now we have that $x \leq f(x)$ and $f(x) \leq x$ which immediately gives $x = f(x)$. $\square$

2.3 Numeric sets and their cardinalities

We now use the axioms of ZFC to construct the most common numeric sets, namely $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$ and $\mathbb{R}$.

Definition 2.10. *Given a set X, we define its successor $s(X) := X \cup \{X\}$.*

Now we set $0 := \emptyset$ and, inductively, $n + 1 := s(n) = \{0, \ldots, n\}$. In particular, for instance, we have $1 = \{\emptyset\}$, then $2 = \{\emptyset, \{\emptyset\}\}$, then $3 = \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$ and so on. We now want to have a set containing all the objects thus defined, i.e. the set of natural numbers. In order to do that, we have to resort to the axiom of infinity.

Definition 2.11. *We will say that a set X is inductive if $\emptyset \in X$ and $y \in X \to s(y) \in X$.*

Observe now that what the axiom of infinity really says is "there exists an inductive set". Now define $\omega := \cap X$ taken over the class of all inductive sets. Note that this is indeed well-defined, because it may be expressed formally as follows. Granted by the axiom of infinity that there is at least one inductive set X, we will use separation to define $\omega = \{x \in X | Y \text{ inductive} \to x \in Y\}$.

The resulting set is sometimes denoted as ω and called the set of finite ordinal numbers, sometimes as $\mathbb{N}$ and called the set of natural numbers. The reason for this double notation will be made clear later on. For the moment, we can just take the two as synonyms for the same object.

By basic abstract algebra, we know that there is an operation of sum on $\mathbb{N}$. We will use that to define integer numbers. The idea is to express an integer as a pair of natural numbers, the first of which can be thought of as having a plus sign in front, and the second as having a minus sign. The formal definition is as follows. Define an equivalence relation on $\mathbb{N} \times \mathbb{N}$ by $(n, m) \sim (n', m')$ if and only if $n + m' = m + n'$ (check that it is indeed an equivalence relation).

Now set $\mathbb{Z} := \mathbb{N} \times \mathbb{N} / \sim$. Observe that whenever $n \geq m$ we have that $(n, m) \sim (n - m, 0)$, and whenever $n < m$ we have that $(n, m) \sim (0, m - n)$. Therefore, the information of an integer amounts to a natural number with the addition of a binary option, that we interpret as sign, precisely as we wanted.

In a very similar way, we want to define rational numbers starting from integers. Again by abstract algebra, we know that there are both an addition and a multiplication on $\mathbb{Z}$. We exploit multiplication in a manner that is much alike what we did above, just being careful not to allow division by zero. So, on the set $\mathbb{Z} \times (\mathbb{Z} \setminus \{0\})$ we define an equivalence relation via $(p, q) \sim (p', q')$ if and only if $pq' = qp'$.

Now set $\mathbb{Q} := \mathbb{Z} \times (\mathbb{Z} \setminus \{0\}) / \sim$.

There are a few ways to define real numbers starting from rationals. We choose the construction through Dedekind cuts, because it turns out that it is the most convenient for what our next purpose.

Definition 2.12. *A subset $A \subseteq \mathbb{Q}$ is called downclosed if $\forall x, y \in \mathbb{Q}$, if $x \in A$ and $y \leq x$ then $y \in A$.*

Definition 2.13. *A Dedekind cut is a proper downclosed subset of $\mathbb{Q}$ that has no maximal element.*

Now we call $\mathbb{R}$ the set of all Dedekind cuts. The idea here is to express each real number as rational approximations from below. We need to requirement that Dedekind cuts be proper subsets in order to prevent $+\infty$ to be an element of $\mathbb{R}$. Furthermore, we need them not to have a maximal element in order to avoid having double representations for all real numbers that happen to be rational.

The next step is, we wonder what the cardinalities of the sets thus defined are. We set by definition $\aleph_0 := |\mathbb{N}|$.

Proposition 2.14. $|\mathbb{Z}| = \aleph_0$.

Proof. By Cantor-Bernstein theorem, it will suffice to define two injective functions $\mathbb{N} \to \mathbb{Z}$ and $\mathbb{Z} \to \mathbb{N}$. The first one can be defined by $n \mapsto +n$.

For the second one, we have to distinguish two cases. Whenever $n \geq 0$, it will be taken to $2n$. If instead $n < 0$, it will be taken to $-2n - 1$.

The two functions are clearly injective, so we are finished. $\square$

Proposition 2.15. $|\mathbb{Q}| = \aleph_0$.

Proof. We use Cantor-Bernstein theorem again. The assignment $n \mapsto \frac{n}{1}$ defines an injective function $\mathbb{N} \to \mathbb{Q}$.

For the opposite direction, we proceed by steps. For each rational number, there exists a unique representation $\frac{p}{q}$ such that p and q are coprime and q is positive. Therefore the assignment $\frac{p}{q} \mapsto (p, q)$ defines an injective function $\mathbb{Q} \to \mathbb{Z} \times \mathbb{Z}$. By the previous proposition, we also have $\mathbb{Z} \times \mathbb{Z} \cong \mathbb{N} \times \mathbb{N}$. We are now reduced to proving that there is an injective function $\mathbb{N} \times \mathbb{N} \to \mathbb{N}$.

We graphically represent the elements of the domain as in the table

(0,0)	(0,1)	(0,2)	(0,3)	$\cdots$
(1,0)	(1,1)	(1,2)	(1,3)	$\cdots$
(2,0)	(2,1)	(2,2)	(2,3)	$\cdots$
(3,0)	(3,1)	(3,2)	(3,3)	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

With this in mind, what we do is sending $(0, 0)$ to 0, then the second ascending diagonal (pairs with sum 1) to the next two naturals, then the third ascending diagonal (pairs with sum 2) to the next three naturals, and so on. More precisely, we define a function by the formula (just for $n + m > 0$)

$$(n, m) \mapsto \left(\sum_{k=1}^{n+m} k \right) + m + 1.$$

It is straightforward to check that this function is indeed injective. Moreover, since for each natural number s there is only a finite number of pairs (n, m) such that $n + m = s$, it will eventually hit all natural numbers, and is thus in fact already surjective, as well. $\square$

Proposition 2.16. $|\mathbb{R}| = |\mathcal{P}(\mathbb{N})|$.

Proof. As above, we use Cantor-Bernstein. The set of real numbers is overtly a subset of $\mathcal{P}(\mathbb{Q})$, so there is an inclusion $\mathbb{R} \subseteq \mathcal{P}(\mathbb{Q}) \cong \mathcal{P}(\mathbb{N})$ by the previous proposition.

For the opposite direction, observe that for a set X, its power set $\mathcal{P}(X)$ is isomorphic to the set of functions $X \to 2$. Such a function can be seen as the subset of X containing precisely all elements that are sent to 1. Now, we define a function $\mathcal{P}(\mathbb{N}) \to \mathbb{R}$ by sending a function $h : \mathbb{N} \to 2$ to the decimal representation $0.h(0)h(1)h(2)\ldots$, which in turn gives a Dedekind cut by successive approximations. This assignment is also easily checked to be injective. $\square$

3 Ordinal numbers and the cumulative hierarchy

We mentioned that there is a double notation for the set of natural numbers, which are sometimes called $\mathbb{N}$, some other times ω. In the context of numeric

sets, the former notation is preferred. The latter is instead used whenever regarding the set of natural numbers as the set of finite ordinals. The ways of extending this set in the two contexts are therefore very different. While we are interested in the algebraic operations defined on natural numbers in order to extend them to integer, rational, real numbers and so on, we will be here interested in their order structure.

3.1 Ordinals

Definition 3.1. *A set X is said to be transitive if $\forall y(y \in X) \to (y \subseteq X)$.*

Definition 3.2. *Let E be a binary relation defined on a set X. We say that X is strictly well-ordered by E if the relation E^+ defined by*

$$xE^+y \leftrightarrow xEy \quad \text{or} \quad x = y$$

gives a well-ordering of X.

Definition 3.3. *An ordinal is a transitive set that is strictly well-ordered by the relation $\in$.*

Remark 3.4. The natural numbers as defined in the previous lectures are ordinals. Moreover, the set ω exhausts all finite ordinals.

We now proceed to define new ordinals that extend the hierarchy we started building with ω. We will do this by using essentially two different tools: the notion of successor and the axiom of replacement.

Given an ordinal α, we can always define its successor $\alpha+1$ by $\alpha \cup \{\alpha\}$, the same way as we did with finite ordinals. Moreover, given a set of ordinals $(\alpha_i)_{i \in I}$, the axiom of union yields that we can define their union $\alpha = \cup_{i \in I} \alpha_i$.

Exercise 3.5. Show that the two procedures above indeed give rise to ordinals.

In particular, iterating the successor construction gives ordinals $\omega + n$ for every finite n. By the axiom of replacement, all the sets of this form form a set, therefore we can again find their union, and we call this $\omega + \omega$. Keeping in mind the same observations, we go on defining:

- $\omega \times n$ as $\omega + \ldots + \omega$, repeated n times;

- $\omega \times \omega$ as the union of all ordinals $\omega \times n$;

- ω^n as $\omega \times \ldots \times \omega$, repeated n times;

- ω^ω as the union of all ordinals ω^n;

- ω^{ω^n} and ω^{ω^ω} similarly;

- $\omega^{\omega^{\cdot^{\cdot^{\cdot^\omega}}}}$ and so on.

Now observe that for an ordinal there are two possibilities: either it is the successor of some other ordinal, or it is not. In the latter case, we say that an ordinal is a limit ordinal. This is the case, for instance, for all the ordinal defined above except for $\omega + n$.

3.2 The cumulative hierarchy and regularity

Using ordinals, we now want to construct a chain of sets such that most constructions can be seen to return sets that are elements of some member of this chain. In fact, assuming the axiom of regularity, one can prove that *all* sets belong to this chain at some point.

- Let us define $V_0 := \emptyset$;

- for an ordinal α, define $V_{\alpha+1} := \mathcal{P}(V_\alpha)$;

- if α is a limit ordinal, then $V_\alpha := \cup_{\beta<\alpha} V_\beta$.

In the third line, we know of course that the union exists because α is a set.

Now we turn our attention to another of the axioms of ZFC that we've been deferring.

- **Axiom of regularity** $\forall x \neq \emptyset \exists y \in x(x \cap y = \emptyset)$

In particular, given an x as in the formula above, such an element y will be referred to as a minimal element of x. Therefore, an alternative formulation of regularity says that every set has a minimal element.

Lemma 3.6. *The following are true:*

1. *if x is a set of ordinals, then $\cup x$ is an ordinal;*

2. *if x is a set of ordinals, then $\exists \alpha$ ordinal such that $x \subseteq \alpha$.*

Proof. Exercise. $\square$

Definition 3.7. *Given a set x, let us define $f_x(0) = x$ and then inductively $f_x(n+1) = \cup f_x(n)$ for all finite ordinals. We will call the transitive closure of x the set $\mathrm{TC}(x) = \cup_{n \in \omega} f_x(n)$.*

Remark 3.8. A set x is transitive if and only if $\mathrm{TC}(x) = x$.

Proposition 3.9. *Assuming the axiom of regularity, then $\forall X \exists \alpha$ such that α is an ordinal and $X \in V_\alpha$.*

In the proof of this theorem, we will use the following abbreviations for a set s:

- $s \in \cup_{\alpha \in \mathrm{ON}} V_\alpha$ for $\exists \alpha$ ordinal such that $s \in V_\alpha$;

- $s \subseteq \cup_{\alpha \in \mathrm{ON}} V_\alpha$ for $\forall t \in s \exists \alpha$ ordinal such that $t \in V_\alpha$.

Proof. Suppose there is a set a that contradicts the statements, that is, $a \notin \cup_{\alpha \in \mathrm{ON}} V_\alpha$. By transitivity of the V_α's, $\mathrm{TC}(a) \notin \cup_{\alpha \in \mathrm{ON}} V_\alpha$. Therefore, without loss of generality, we may assume that a is transitive.
Suppose then that $a \subseteq \cup_{\alpha \in \mathrm{ON}} V_\alpha$. For all $x \in a$, let us say that the rank of x is the least α such that $x \subseteq V_\alpha$. The range of the rank function on a is a set of ordinals, which is by Lemma 3.6(2) a subset of some ordinal β. Hence $a \subseteq V_\beta$, so $a \in V_{\beta+1}$, which is a contradiction. Therefore we may assume $a \not\subseteq \cup_{\alpha \in \mathrm{ON}} V_\alpha$. Hence there must be some $x \in a$ with $x \notin \cup_{\alpha \in \mathrm{ON}} V_\alpha$. Let $b = \{x \in a | x \notin$

$\cup_{\alpha \in \mathrm{ON}} V_\alpha\}$. Since $b \neq \emptyset$, it has a minimal element y. Since $y \notin \cup_{\alpha \in \mathrm{ON}} V_\alpha$, we can argue as above that $y \nsubseteq \cup_{\alpha \in \mathrm{ON}} V_\alpha$, so there is z such that $z \in y$ and $z \notin \cup_{\alpha \in \mathrm{ON}} V_\alpha$. Since a is transitive, we also have that $z \in a$ and therefore, by the definition of b, $z \in b$. We conclude that $z \in y \cap b$, which is a contradiction because y was taken to be minimal in b. $\qquad \square$

Definition 3.10. *Given a set X, we say that X has rank α, and write $\mathrm{rk}\, X = \alpha$, if $\alpha = \min\{\alpha | X \in V_\alpha\}$.*

Remark 3.11. It is clear that assuming regularity, it follows from Proposition 3.9 that every set has a well defined rank.

Theorem 3.12. *Regularity holds if and only if every set is an element of some V_α.*

Proof. One direction has already been proven in Proposition 3.9. For the converse, suppose that $x \neq \emptyset$ is a set and find a minimal element of it. By assumption and the remark above, we may define a rank function on x, whose range is a set a. Let α be minimal in a, and let $y \in x$ such that $\mathrm{rk}\, y = \alpha$. Then we must have $y \cap x = \emptyset$, because otherwise we would have $z \in y \cap x$, therefore $\mathrm{rk}\, z < \mathrm{rk}\, y$, which contradicts the minimality of α. $\qquad \square$

Lemma 3.13. *The following statements are true:*

1. *if $\alpha < \beta$ are ordinals, then $V_\alpha \subset V_\beta$;*

2. *if $\alpha < \beta$ are ordinals, then $V_\alpha \in V_\beta$;*

3. *if α is an ordinal, then $\mathrm{rk}\, \alpha = \alpha$;*

4. *if X is a set and $\exists \max(\mathrm{rk}\, x)_{x \in X}$, then $\mathrm{rk}\, X = \max(\mathrm{rk}\, x)_{x \in X} + 1$;*

5. *if X is a set and $\nexists \max(\mathrm{rk}\, x)_{x \in X}$, then $\mathrm{rk}\, X = \sup(\mathrm{rk}\, x)_{x \in X}$;*

6. *$\mathrm{rk}\, (a, b) = \max\{\mathrm{rk}\, a, \mathrm{rk}\, b\} + 2$.*