These are the notes to the lecture on ZFC as a foundation of mathematics given as part of the examination on the course "Types and programming languages" (the formal PhD-course in the Initial Types Club) at Chalmers university of technology and Gothenburg university in spring 2018.

The material in these notes is mainly taken from [2], [3] and [1]. Since the it is mainly intended to be orally presented, these notes are somewhat sketchy. In particular, the treatment of parentheses will be sloppy.

Outline

Two topics on ZFC as a foundational theory of mathematics:

- Construction of ordered pairs, Cartesian products, sets of functions and quotient sets.
- Construction of the natural numbers.

Contents

1 The language of ZFC

We consider ZFC as a theory in first-order logic with equality. The language of ZFC has only one non-logical symbol, the binary relation symbol $\in$. As such, the language contains no closed or complex terms.

We will however use defined concepts in this language, many of which will be usable like closed terms in that we can prove in ZFC that they are satisfied by a unique object. The (probably) simplest example would be that of the empty set, which we define as

$$[x = \varnothing] \text{ is an abbreviation of the formula } \forall y(y \notin x)$$

(where of course x and y are distinct variables). It can be shown that

$$\mathrm{ZFC} \vdash \exists x[x = \varnothing] \wedge \forall y([y = \varnothing] \to y = x).$$

We therefore informally use the symbol $\varnothing$ as a constant of the language. The same goes for many other common notions (like $\subseteq$, $\times$ and $\mathbb{N}$) (though there is a small issue here, since $\varnothing$ and $\subseteq$ are in some sense purely set-theoretic concepts; thus it does not matter which is the actual longhand of e.g. $[x \subseteq y]$, since they can all be shown (in ZFC) to be equivalent or define the same sets. This is not so for $\times$ and $\mathbb{N}$; it can be shown that there are several different formalisations of these concepts). However, as long as we only use only the properties of these objects which they are supposed to formalise (see e.g. the definitions of Cartesian products and $\mathbb{N}$), the propositions are equivalent. If we include "lower-level" concepts though, we can get some silly results like $2 \in \pi$. While most of these constructions will turn out to be unique up to isomorphism, what this means for basic concepts like that of function itself is a bit unclear.

In addition to "being the empty set" above, we will (non-exclusively) use the following abbreviations in the axioms which follow:

$$\begin{aligned}
[\varnothing \in x] &:= \forall z(z = \varnothing \to z \in x) \\
[\forall x \in y\, \varphi] &:= \forall x(x \in y \to \varphi) \\
[x \subseteq y] &:= \forall z \in x(z \in y) \\
[x \subset y] &:= x \subseteq y \wedge x \neq y
\end{aligned}$$

We will treat other class-functions (such as $\mathcal{P}$; see below) like $\varnothing$ (when writing e.g. $\mathcal{P}(x) \in y$). As already mentioned (and used) we will also be sloppy with parentheses and brackets, writing them where it seems most clarifying. The same goes for variable names.

While we will mainly take an external perspective on ZFC, writing ZFC $\vdash \varphi$ when stating that φ is a theorem of ZFC, when doing proofs and similar reasoning (and later also when presenting more complex theorems) we will simplify the process taking an internal perspective; quantifiers etc. then ranges over the universe of sets.

1.1 Classes and class-functions

Instead of manipulating formulae one often introduces classes. The class corresponding to a formula is thought of as "the collection of objects" satisfying the formula. One then reserves the braces-notation for classes and writes $\{x \,|\, \varphi\}$ for the class corresponding to φ, with the convention that certain classes also are sets, as given by the axioms. However, I personally find this practice slightly confusing and will not use it in the following.

A more useful notion is that of a class-function. A class function (as we will use the term) is a formula which ZFC proves to have functional properties, that is $\varphi(x, y)$ is a class-function if ZFC $\vdash \forall \bar{x} \exists y(\varphi(\bar{x}, y) \wedge \forall z(\varphi(x, z) \to z = y))$. We often abbreviate such formulae with a more suggestive notation, treating the abbreviations as function symbols (and constants) of the language (like the example of $\varnothing$ earlier). Also, like all functions, class-functions can be composed to yield new class-functions (e.g. with the definition $\exists z(\psi(x, z) \wedge \varphi(z, y))$).

2 The axioms

The following formulae and formula schemata constitute the axioms of ZFC.

Extensionality $\forall x, y(x = y \leftrightarrow \forall z(z \in x \leftrightarrow z \in y))$. Two sets are equal if and only if they have the same members.

Pairing $\forall x, y \exists p \forall z(z \in p \leftrightarrow z = x \vee z = y)$. For any two sets their pair is a set (and hence each singleton is a set). Note that by extensionality this set is unique, whence $\forall z(z \in p \leftrightarrow z = x \vee z = y)$ is a class-function. We abbreviate this as $[p = \{x, y\}]$ (or $[p = \{x\}]$ if $x = y$).

Separation (Subsets) For every formula $\varphi(z, \bar{p})$: $\forall \bar{p}, A \exists B \forall z(z \in B \leftrightarrow z \in A \wedge \varphi(z, \bar{p}))$. "All"[1] subsets (which can be defined) exists. Once again, uniqueness follows from extensionality and $\forall z(z \in B \leftrightarrow z \in A \wedge \varphi(z, \bar{p}))$ is a class-function (in $A, \bar{p}$). We write $[B = \{z \in A \,|\, \varphi(z, \bar{p})\}]$.

Union $\forall A \exists x \forall z(z \in x \leftrightarrow \exists y \in A(z \in y))$. For every set (of sets) there is a set consisting of the union of its elements. This is also a class-function which we denote $[x = \bigcup A]$. We also write $[x = u \cup v]$ if $A = \{u, v\}$ (from pairing).

Power Set $\forall x \exists P \forall z(z \in P \leftrightarrow z \subseteq x)$. For every set there is a "power set" consisting of its subsets. We denote this class-function $[P = \mathcal{P}(x)]$.

Infinity $\exists I(\varnothing \in I \wedge \forall z \in I(z \cup \{z\} \in I))$. This axiom states that there is a set which is *inductive* (and in particular at least one set (note that the existence of $\varnothing$ is not required for this to make sense)). We write $\mathrm{Ind}(x)$ (NB: this is not a class-function, but a defined predicate).

Replacement For every formula $\varphi(x, z, \bar{p})$: $\forall \bar{p}(\forall x, y, z(\varphi(x, z, \bar{p}) \wedge \varphi(x, y, \bar{p}) \to y = z) \to \forall A \exists B \forall z(z \in B \leftrightarrow \exists x \in A \varphi(x, z, \bar{p})))$. Essentially, the image of a set under a class-function is a set (i.e. every class-function restricted to a set is actually a set function), except we have defined class functions to be total, and this is stated "for those values of the parameters" which the formula defines a total function.

Foundation (Regularity) $\forall x(x \neq \varnothing \to \exists z \in x(z \cap x = \varnothing))$. This means that the membership-relation $\in$ is well-founded. This is in principle not used for ordinary mathematics.

Choice $\forall A(\varnothing \notin A \wedge \forall x, y \in A(x \cap y = \varnothing) \to \exists S(S \subseteq \bigcup A \wedge \forall x \in A \exists y \in x(x \cap S = \{y\})))$. This is most often given in some more readable formulation like $\forall A, B \forall r \subseteq A \times B(\forall x \in A \exists y \in B(r(x, y)) \to \exists f : A \longrightarrow B \forall x \in A(r(x, f(A))))$. Relative the other axioms these sentences are equivalent.

[1] See Skolem's paradox.

3 Basic constructions

3.1 Cartesian products

First we want a notion of ordered pair (as opposed to the unordered pair $\{x, y\}$). That is, we want a class-function $[z = (x, y)]$ (with the displayed free variables) such that

$$\mathrm{ZFC} \vdash \forall x, y, z, u, v(([z = (x, y)] \wedge [z = (u, v)]) \to (u = x \wedge v = y)) \tag{1}$$

Of course, the smaller the fragment of ZFC we need to prove this, the better (to some extent). In order to be able to construct Cartesian products of arbitrary sets we need also require that there is some set containing all ordered pairs from the two sets, that is:

$$\mathrm{ZFC} \vdash \forall A, B \exists C(\forall x \in A, y \in B((x, y) \in C)) \tag{2}$$

We prove there is such a formula and then forget the particular choice; thenceforth we treat $(\cdot, \cdot)$ as a function symbol of the language and only use properties (1) and (2).

The (as far as I know) most common choice is the Kuratowski pair:

$$[z = (x, y)] := z = \{\{x\}, \{x, y\}\}.$$

We show that it satisfies the above properties.

Proof. [2] By the pairing and extensionality axioms $[z = \{x, y\}]$ and $[z = \{x\}]$ are class-functions, whence the Kuratowski pairing formula is as well.

To prove (1) take x, y, u and v. Assume $(x, y) = (u, v)$. Then, since $\{x\} \in (x, y)$, $\{x\} \in (u, v)$. Thus $\{x\} = \{u\}$ or $\{x\} = \{u, v\}$. In either case $u \in \{x\}$ (since $u \in \{u\}$ and $u \in \{u, v\}$), whence $u = x$. Moreover, $\{x, y\} \in (x, y)$ whence $\{x, y\} = \{u\}$ or $\{x, y\} = \{u, v\}$. Thus we have $y = u \vee (y = u \vee y = v)$. If $y = u = x$ then $(x, y) = \{\{x\}\}$, i.e. $\{\{x\}\} = \{\{u\}, \{u, v\}\}$, whence $\{x\} = \{u, v\}$ and $v = x = y$. In either case $v = y$.

For (2), take A and B. Since for $x \in A$, $y \in B$ the sets $\{x\}, \{x, y\} \subseteq A \cup B$ we have $\{x\}, \{x, y\} \in \mathcal{P}(A \cup B)$. Thus $\{\{x\}, \{x, y\}\} \in \mathcal{P}(\mathcal{P}(A \cup B))$. □

We can now define the Cartesian product by the class-function $[C = A \times B] := [C = \{z \in \mathcal{P}(\mathcal{P}(A \cup B)) \mid \exists x \in A \exists y \in B(z = (x, y))\}]$ by separation (as always, uniqueness follows by extensionality).

n-tuples and n-fold Cartesian products are defined by iteration, e.g. $(x_1, \ldots, x_n, x_{n+1}) = ((x_1, \ldots, x_n), x_{n+1})$ and $A_1 \times \cdots \times A_n \times A_{n+1} = (A_1 \times \cdots A_n) \times A_{n+1}$ (though I am not aware of any consensus regarding whether the iteration should be to the left or right). Note that these constitute one class-function *for each* n; the n is not an argument to a single function.

Also, it follows by foundation that $x \neq (x, y) \neq y$ for all x, y.

[2] This proof is adapted from [1]. Note that the axioms are only required to show that the pair exists and is unique (i.e. that the pairing formula is a class function), the rest is pure (and constructive/intuitionistic) logic.

3.2 Relations and functions

We identify relations and functions with their graphs. That is

$$\begin{aligned}\mathrm{Rel}_2(r, A, B) &:= r \subseteq A \times B\\ \mathrm{Rel}_n(r, A_1, \ldots, A_n) &:= r \subseteq A_1 \times \cdots A_n\end{aligned}$$

and similarly for functions

$$f : A \longrightarrow B := \mathrm{Rel}_2(f, A, B) \wedge \forall x \in A \exists y \in B((x, y) \in f \wedge \forall z \in B((x, z) \in f \to z = y)).$$

Moreover, we define

$$\begin{aligned}R(x_1, \ldots, x_n) &:= (x_1, \ldots, x_n) \in R\\ [f(x) = y] &:= (x, y) \in f\end{aligned}$$

(though we only use this notation for relations and functions, respectively).

With this definition of function, the set of functions between A and B is given by the class-function

$$[F = {}^A B] := F = \{f \in \mathcal{P}(A \times B) \mid f : A \longrightarrow B\}.$$

For any A and B, this set exists by separation and is unique by extensionality. Other common notations for this are B^A and $(A \longrightarrow B)$. Injections, surjections and bijections are then defined as usual.

NB: The codomain is not determined by the function; if $f : A \longrightarrow B$ and $B \subseteq C$, then $f : A \longrightarrow C$ as well. This has some advantages, for instance when constructing functions, but also has disadvantages (e.g. when doing category theory, morphisms in the category of sets must be something more than just the functions (for instance the pair of the function and its intended codomain)).

3.3 Quotient sets

In ZFC we can also construct quotient sets of equivalence relations as sets of equivalence classes. Thus we have a binary class function $[x = A/\!\sim]$ such that, if $\mathrm{Eq}(\sim, A)$ is a formula stating that $\sim$ is an equivalence relation on A,

$$\begin{aligned}\mathrm{ZFC} \vdash \forall A, \sim (\mathrm{Eq}(\sim, A) \to (&\forall x \in A/\!\sim (x \neq \varnothing) \wedge\\ &\left(\bigcup A/\!\sim\right) = A \wedge \forall x, y \in A/\!\sim (x \neq y \to x \cap y = \varnothing)\\ &\wedge \forall a, b \in A(a \sim b \leftrightarrow \exists x \in A/\!\sim (a \in x \wedge b \in x))).\end{aligned}$$

This is given via the definitions

$$\begin{aligned}[x = [a]_A] &:= x = \{b \in A \mid b \sim a\},\\ [Q = A/\!\sim] &:= Q = \{y \in \mathcal{P}(A) \mid \exists a \in A(y = [a]_A)\}.\end{aligned}$$

Note in particular that a consequence of this construction is that the equality on the quotient set is the ordinary equality.

4 The natural numbers

The natural numbers should be a triple $(N, 0, S)$ satisfying the Peano axioms[3]

$$
\begin{aligned}
&0 \in N,\\
&S : N \longrightarrow N,\\
&\forall x, y \in N(S(x) = S(y) \to x = y),\\
&\forall x \in N(S(n) \neq 0),\\
&\forall A \in \mathcal{P}(N)(0 \in A \wedge \forall x \in A(S(x) \in A) \to A = N).
\end{aligned}
$$

These properties characterise the set of natural numbers up to isomorphism in ZFC[4]. The usual (and in some sense most natural) choice is to identify the natural numbers with the finite ordinal numbers and the set of naturals with the first infinite ordinal. We thus turn briefly to ordinal numbers (which are important in themselves in set theory).

4.1 Ordinal numbers

We define transitive sets as follows

$$\mathrm{Trans}(x) := \forall y \in x \forall z \in y(z \in x).$$

Colloquially put, $z \in y \in x \to z \in x$. An ordinal is a transitive set well-ordered by $\in$:

$$
\begin{aligned}
\mathrm{Ord}(\alpha) := \mathrm{Trans}(\alpha) \wedge \forall x, y \in \alpha(x \notin x \wedge (x \in y \vee y \in x \vee x = y))\\
\wedge \forall A \in \mathcal{P}(\alpha)(A \neq \varnothing \to \exists x \in A \forall y \in A(x \in y \vee x = y))
\end{aligned}
$$

Note that only transitivity and linearity are really required of an ordinal; both irreflexivity and well-foundedness then follows by the axiom of foundation.

The ordinals are "well-ordered" by $\in$ and the successor is given by the class function

$$[x = \alpha^+] := x = \alpha \cup \{\alpha\}.$$

The successor of an ordinal is again an ordinal. $\varnothing$ is an ordinal and the smallest such. We will also use the following straightforward facts about ordinals:

1. Every element of an ordinal is an ordinal.
2. If $\alpha \subset \beta$ are ordinals, then $\alpha \in \beta$.

Recall that a set is inductive if it contains $\varnothing$ and is closed under the successor operation, and that the axiom of infinity postulates the existence of an inductive set. It thus seems intuitive that the smallest inductive set is an ordinal which in addition satisfies the Peano axioms above. We thus define the set ω by the class-function

$$[x = \omega] := \forall z(z \in x \leftrightarrow \forall A(\mathrm{Ind}(A) \to z \in A)).$$

We show that this is indeed a class function (i.e. that ω exists), that ω is an inductive ordinal and that $(\omega, \varnothing, \_^+)$ satisfies the Peano axioms.

[3] See [5]
[4] See [3].

Proof. Let I be some inductive set, the existence of which is guaranteed by the axiom of infinity. Let $a = \{x \in I \,|\, \forall A(\mathrm{Ind}(A) \to x \in A)\}$ which exists by the axiom of separation. Then any x such that $\forall A(\mathrm{Ind}(A) \to x \in A)$ in particular satisfies $x \in I$, whence $x \in a$. Conversely, $x \in a$ implies $\forall A(\mathrm{Ind}(A) \to x \in A)$ by definition. Hence $a = \omega$. Extensionality shows that a is unique with this property.

Since $\varnothing \in A$ for all inductive A, $\varnothing \in \omega$. Moreover, if $x \in \omega$ then $x \in A$ for all inductive A, whence $x^+ \in A$ for all such A by inductivity and hence $x^+ \in \omega$. So $\mathrm{Ind}(\omega)$.

This also verifies the first two Peano axioms, and the fourth is immediate. To verify the fifth, let $A \subseteq \omega$ be such that $\varnothing \in A$ and $x \in A$ implies $x^+ \in A$. Then A is inductive whence $\omega \subseteq A$ by definition, so $A = \omega$.

Now let $B = \{x \in \omega \,|\, x \subseteq \omega\}$. Then $\varnothing \in B$ trivially, and if $y \in B$ then by definition $y \in \omega$ and $y \subseteq \omega$, whence $y^+ = y \cup \{y\} \subseteq \omega$. Hence $B = \omega$ which means that every element of ω is a subset of ω. This is equivalent to transitivity.

Let $C = \{x \in \omega \,|\, \mathrm{Ord}(x)\}$. Since $\varnothing \in C$ and if $x \in C$ then $x^+ \in C$ by above, $C = \omega$. Since the ordinals are linearly ordered by $\in$ and every non-empty set of ordinals has a $\in$-least element, it follows that ω is well-ordered by $\in$.

Thus ω is an ordinal number. It now only remains to verify the third Peano axiom. To this end, let $x, y \in \omega$ and suppose $x^+ = y^+$, that is $x \cup \{x\} = y \cup \{y\}$. Then $x \in y$ or $x = y$, and $y \in x$ or $y = x$. That $x \in y$ and $y \in x$ contradicts the linearity of $\in$ in ω (and the axiom of foundation), and in all other cases $x = y$. This concludes the proof. □

This is not the easiest construction of $\mathbb{N}$, but as ω is an important object in set theory itself, it seems reasonable to illustrate this. However, similarly to the case of ordered pairs we need not (and perhaps should not) care which particular triple is chosen for $\mathbb{N}$, as long as it satisfies the Peano axioms. For example, while addition, multiplication and the ordering on $\mathbb{N}$ coincides with the particular ones for ω viewed as an ordinal, they can also be defined by recursion.

Theorem 1 (Recursion theorem)**.** $\mathrm{ZFC} \vdash \forall A \forall a \in A \forall g : A \longrightarrow A \exists f : N \longrightarrow A(f(0) = a \wedge f(S(n)) = g(f(n)))$.

Proof. Let

$$\begin{aligned} D =& \{i \in \mathcal{P}(N) \,|\, 0 \in\ i \wedge \forall n \in N(S(n) \in i \to n \in i)\} \text{ and} \\ F =& \{p \in \mathcal{P}(N \times A) \,|\, \exists i \in D \\ & \qquad (p : i \longrightarrow A \wedge p(0) = a \wedge \forall n \in i(S(n) \in i \to p(S(n)) = g(p(n))))\}. \end{aligned}$$

Take $i, j \in D$ and $p, q \in F$ with $p : i \longrightarrow A$ and $q : j \longrightarrow A$. Let $X = \{n \in N \,|\, n \in i \cap j \to p(n) = q(n)\}$. By definition of F, $p(0) = a = q(0)$ and thereby $0 \in X$. Moreover, if $n \in X$ and $S(n) \in i \cap j$, then $n \in i \cap j$ by definition of D, so $p(n) = q(n)$ by the definition of X and hence $p(S(n)) = g(p(n)) = g(q(n)) = q(S(n))$ by the definition of F. Hence $X = N$ by induction whence $n \in i \cap j$ entails $p(n) = q(n)$ for all $n \in N$. So any two functions from F agree where they are both defined.

Now let $p \in F$ and $i \in d$ be such that $p : i \longrightarrow N$. Suppose $n \in i$ and consider $p' = p \cup \{(S(n), g(p(n)))\}$ and $i' = i \cup \{S(n)\}$. Clearly $i' \in D$ and $p' \subseteq i' \times A$. If

$S(n) \in i$ then $i' = i$ and $p(S(n)) = g(p(n))$, whence $(S(n), g(p(n))) \in p$ and $p' = p$. Otherwise, if $(m, y), (m, z) \in p'$ then either $(m, y), (m, z) \in p$, whence $y = z$, or $(m, y) \in p$ and $(m, z) = (S(n), g(b))$ (or the other way around) which contradicts $S(n) \notin i$, or $(m, y) = (S(n), g(b)) = (m, z)$, whence $y = z$. Hence $p' : i' \longrightarrow A$. In either case $(0, a) \in p \subseteq p'$, and if $m \in i'$ is such that $S(m) \in i'$ then either $S(m) \in i$ and $p'(S(m)) = p(S(m)) = g(p(m)) = g(p'(m))$ or $m = n$ and $p'(S(m)) = g(p(m))$ by definition. Hence $p' \in F$. So the members of F can be extended with a "next" value.

Define $f = \bigcup F$. Then if $x \in f$ there is some $p \in F$ with $x \in p$, whence $x \in N \times A$. So $f \subseteq N \times A$ is a relation. Let $d = \{n \in N \mid \exists b \in A((n, b) \in f)\}$. Since $\{0\} \in D$ and hence $\{(0, a)\} \in F$ we have $(0, a) \in f$ and $0 \in d$. Moreover, suppose $n \in d$. Then $(n, b) \in f$ for some $b \in A$, that is $n \in i$ and $p(n) = b$ for some $i \in D$ and $p \in F$ with $p : i \longrightarrow A$. By above there is a $p' \in F$ such that $(S(n), g(b)) \in p'$, whence $(S(n), g(b)) \in f$ and $S(n) \in d$. Thus $d = N$.

Now let $(n, b), (n, c) \in f$. Then $p(n) = b$ and $q(n) = c$ for some $p, q \in F$. If $i, j \in D$ are such that $p : i \longrightarrow A$ and $q : j \longrightarrow A$, then $n \in i \cap j$ and $b = p(n) = q(n) = c$ by the firs paragraph. Hence $f : N \longrightarrow A$. As noted in the previous paragraph, $f(0) = a$. Also, for all $n \in N$ there is a $p \in F$ such that $p(n) = f(n)$, whence there is a $p' \in F$ with $p'(S(n)) = g(p(n))$ and $f(S(n)) = p'(S(n)) = g(f(n))$. Thus f satisfies the recursive equations and in particular (since $N \in D$) $f \in F$. If $h : N \longrightarrow A$ also satisfies the recursive equations then $h \in F$ as well, whence $f(n) = g(n)$ for all $n \in N$ by the first paragraph. □

References

[1] Peter Aczel and Michael Rathjen. Notes on constructive set theory. Technical Report 40 2000/2001, Institut Mittag-Leffler, July 2001. The Royal Swedish Academy of Sciences; ISSN 1103-467X; ISNR IML-R- -40-00/01- -SE.

[2] Thomas J. Jech. *Set Theory*. Springer Monographs in Mathematics. Springer, third millennium edition, 2002.

[3] Yiannis Moschovakis. *Notes on Set Theory*. Undergraduate Texts in Mathematics. Springer, second edition, 2006.

[4] Guiseppe Peano. *Arithmetices principia, nova methodo exposita*. Bocca and Clausen, 1889. English translation of the relevant part: [5].

[5] Guiseppe Peano. The principels of arithmetic, presented by a new method. In Jean van Heijenoort, editor, *From Frege to Gödel: A Source Book in Mathematical Logic*, Source Books in the History of the Sciences, pages 85 – 97. Harvard University Press, 1967. English translation of central parts of [4].