Chapter 1. Informal introdution to the axioms of ZF.*

1.1. Extension. Our conception of sets comes from set of objects that we know well such as $\mathbb{N}$, $\mathbb{Q}$ and $\mathbb{R}$, and subsets we can form from these determined by their properties. Here are two very simple examples:

$$\{ r \in \mathbb{R} \mid 0 \le r \} = \{ r \in \mathbb{R} \mid r \text{ has a square root in } \mathbb{R} \}$$

$$\{ x \mid x \neq x \} = \{ n \in \mathbb{N} \mid n \text{ even, greater than two and a prime} \}.$$

(This last example gives two definitions of the *empty set*, $\emptyset$).

The notion of set is an abstraction of the notion of a property. So if X, Y are sets we have that $\big(\forall x\,(\,x \in X \iff x \in Y)\big) \implies X = Y$.

One can immediately see that this implies that the empty set is unique.

1.2. Unions. Once we have some sets to play with we can consider sets of sets, for example: $\{ x_i \mid i \in I \}$ for I some set of indices. We also have the set that consists of all objects that belong to some x_i for $i \in I$. The indexing set I is not relevant for us, so $\bigcup\{ x_i \mid i \in I \} = \bigcup_{i \in I} x_i$ can be written $\bigcup X$, where $X = \{ x_i \mid i \in I \}$.

Typically we will use a, b, c, w, x, z, $\ldots$ as names of sets, so we can write

$$\bigcup x = \bigcup \{ z \mid z \in x \} = \{ w \mid \exists z \,\text{such that}\, w \in z \text{ and } z \in x \}.$$

Or simply: $\bigcup x = \{ w \mid \exists z \in x \; w \in z \}$.

(To explain this again we have

$$\bigcup_{y \in x} y = \bigcup \{ y \mid y \in x \} = \bigcup \{ z \mid \exists y \in x \; z \in y \}.$$

So we have already changed our way of thinking about sets and elements as different things to thinking about all sets as things of the same sort, where the elements of a set are simply other sets.)

1.3. Power sets If x is a set then $\mathcal{P}(X) = \{ Z \mid Z \subseteq X \}$, the set of all subsets of x is also a set.

Proposition. (Cantor) There is no surjection $f : X \longrightarrow \mathcal{P}(X)$ (a function $f : X \longrightarrow \mathcal{P}(X)$ such that for all $Z \in \mathcal{P}(X)$ there is $x \in X$ such that $f(x) = Z$.

Proof. If f was such a surjection we could consider $\{ x \in X \mid x \notin f(x) \}, = A$, let's say. As $A \subseteq X$ we have that $A \in \mathcal{P}(X)$ and so there is some $a \in X$ such that $A = f(a)$ since f is a surjection. But look, now we have that $a \in A$ if and only if $a \notin A$, uma contradiction!

Corollary. The collection of all sets is not a set.

Proof. Let's call the collection of all sets V. If V was a set then $\mathcal{P}(V)$ would be a set as well. But $\mathcal{P}(V)$ would be the set of all subsets of V, so for all $x \in \mathcal{P}(V)$ we would have $x \in V$ and hence $\mathcal{P}(V) \subseteq V$. But in that case there would be a surjection $f : V \longrightarrow \mathcal{P}(V)$ given by $x \mapsto x$ if $x \in \mathcal{P}(V)$ and $x \mapsto \emptyset$ if $x \notin \mathcal{P}(V)$. Contradiction!

* ver 1.1, 12/2/06

Propoganda slogan: "There is no universal set!"

This corollary is sometimes called *Cantor's paradox*. *Russell's paradox* can be obtained in a similar way by taking f to be the identity function in the proof of the proposition above to see that there can be no collection $\{x \mid x \notin x\}$. (That is, if $\{x \mid x \notin x\}, = r$, let's say, was a set, then the question of whether $r \in r$ or $r \notin r$ would give a similar contradiction.) Clearly these are not true paradoxes, but rather memorably/impressive facts.

"Russell's paradox" shows that we cannot hope to use the extensions of all mathematical properties as sets – some are inherently contradictory – and this gives us one reason to need to develop a satisfactory theory of sets. However I really ought to say that set theory began and continues to be interesting and useful **not** to avoid paradoxes or to give secure foundations to mathematics, **but** as a part of mathematics (just like its other parts) with its own theory, questions, techinques, etc., and mathematical links with other areas of mathematics.

1.4. The cummulative hierarchy. We will see what we get if we start with the empty set successively take power sets and unions. We begin with $V_0 = \emptyset$ and repeatedly apply $\mathcal{P}$.

$$\begin{aligned} V_0 &= \emptyset \\ V_1 &= \mathcal{P}(V_0) = \{\emptyset\} \\ V_2 &= \mathcal{P}(V_1) = \{\emptyset, \{\emptyset\}\} \\ V_3 &= \mathcal{P}(V_2) = \{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \{\emptyset, \{\emptyset\}\}\} \\ &\vdots \end{aligned}$$

After all of the finite steps we collect together what we have so far and get: $V_\omega = \bigcup\{V_n \mid n \text{ finite}\}$.

Now we can begin to apply $\mathcal{P}$ again: $V_{\omega+1} = \mathcal{P}(V_\omega)$, $V_{\omega+2} = \mathcal{P}(V_{\omega+1})$, ...

One can continue in this way "generating" sets in one's imagination "for ever."

(The "V"s are so-named for *von Neumann*, the first person to think about this hierarchy.)

Definition. A *set* is a member of some V_α.

The *universe* of sets (generated from $\emptyset$) is the collection

$$V = \{x \mid \exists\alpha\,(\alpha \text{ indexes a step in the iteration} \,\&\, x \in V_\alpha)\}.$$

As a shorthand we can talk about $V =$ "$\bigcup$"$\{V_\alpha \mid \alpha \text{ indexes a step in the iteration}\}$, but as we have already seen, V is not a set and this "union" is not the union of a set – we have that each V_α is a set but we cannot properly take the union of them and form V *as a set*. So we have to think of V in some sense as an uncompleted entity, and hence, V really is just an abbreviation for us – we'll only use it (later) in occasions such as "$x \in V$" by which we formally will mean "there is an α such that $x \in V_\alpha$." Nevertheless it is useful to have a name in hand for this collection of all sets.

Proposition. If α comes before β as an index then $V_\alpha \subseteq V_\beta$.

Proof. This is clear if $\alpha = \beta$. It is also clear if V_β was generated by the process of taking a union (because $V_\alpha \in \{V_\gamma \mid \gamma \text{ is before } \beta \text{ as an index}\}$ and so if $x \in V_\alpha$ we have that

$$x \in \bigcup\{V_\gamma \mid \gamma \text{ is before } \beta \text{ as an index}\}).$$

However, if $V_\beta = \mathcal{P}(V_\gamma)$ for some γ we have to show that $V_\alpha \subseteq V_\gamma \implies V_\alpha \subseteq \mathcal{P}(V_\gamma)$ in order to complete the proof.

So let me interupt the proof to give an important definition.

1.5: Transitive sets.

Definition. A set x is *transitive* if $u \in v \in x \Longrightarrow u \in x$ (if and only if $u \in x \Longrightarrow u \subseteq x$, if and only if $\bigcup x \subseteq x$).

To finish the proof it suffices to show that each V_α is transitive. Because, as $V_\alpha \subseteq V_\gamma$ we have $x \in V_\alpha \Longrightarrow x \in V_\gamma$, and if V_γ is transitive we have $x \in V_\gamma \Longrightarrow x \subseteq V_\gamma$. Thus $x \in V_\alpha \Longrightarrow x \subseteq V_\gamma$, so $x \in \mathcal{P}(V_\gamma) = V_\beta$.

But as the following proposition shows, all V_α are transitive.

Proposition. (i) $\emptyset$ is transitive. (ii) If x is transitive then $\mathcal{P}(x)$ is transitive. (iii) Unions of transitive sets are also transitive.

Proof. (i) and (iii) are both trivial. (Can you see why?) (ii) If $u \in v \in \mathcal{P}(x)$, we have $u \in v \subseteq x$ and thus $u \in x$. But since x is transitive we have that $u \subseteq x$. Hence $u \in \mathcal{P}(x)$.

NOTE! To say that x is transitive is not the same thing as saying that $\in$ is a transitive relation on x. ($\in$ is a transitive relation on x if every element of x is transitive.)

We have implicitly used the principle of induction that

$$\forall\alpha\,\big((\forall\beta < \alpha\,\mathrm{Pr}(\beta)) \longrightarrow \mathrm{Pr}(\alpha)\big) \longrightarrow \forall\alpha\,\mathrm{Pr}(\alpha),$$

where Pr(.) is a property of our indices. We will use this principle to show the following:

Proposition. ($\in$-induction) $\forall x\,\big((\forall u \in x\,\phi(u)) \longrightarrow \phi(x)\big) \longrightarrow \forall x\,\phi(x)$, for any property ϕ.

Proof. We show by induction on α that $\forall x \in V_\alpha\,\phi(x)$. Let us give this property of α, $\forall x \in V_\alpha\,\phi(x)$, the name $\mathrm{Pr}(\alpha)$. By our induction principle it is sufficient to show for each α that we have $\forall\beta < \alpha\,\mathrm{Pr}(\beta)) \longrightarrow \mathrm{Pr}(\alpha)$.

So, suppose that $\forall\beta < \alpha\,\mathrm{Pr}(\beta)$. Take $x \in V_\alpha$. We have two cases.

Either $V_\alpha = \mathcal{P}(V_\beta)$ for some β less that α, when $x \subseteq V_\beta$ and for all $u \in x$ we have $\phi(u)$. Then using the antecedente for x we have $\phi(x)$. Because x was arbitrary we have $\forall x \in V_\alpha\,\phi(x)$, *i.e.*, $\mathrm{Pr}(\alpha)$.

Or $V_\alpha = \bigcup_{\beta<\alpha} V_\beta$, when there is some $\beta < \alpha$ such that $x \in V_\beta$. As we already know $\mathrm{Pr}(\beta)$ we have $\phi(x)$. Again, because x was arbitrary we have $\forall x \in V_\alpha\,\phi(x)$, *i.e.*, $\mathrm{Pr}(\alpha)$.

As α was arbitrary we have $\forall\alpha\,\big((\forall\beta < \alpha\,\mathrm{Pr}(\beta)) \longrightarrow \mathrm{Pr}(\alpha)\big)$. And now we can deduce that $\forall\alpha\,\mathrm{Pr}(\alpha)$, and hence $\forall\alpha\,\forall x \in V_\alpha\,\phi(x)$, and thus $\forall x\,\phi(x)$.

In order to avoid doing too many proofs like this we use a principle, Foundation, equivalent to $\in$-induction:

($\star$) If A is a property of sets and there is some x such that $A(x)$ then there is some x such that $A(x)$ and for all $u \in x$ we do not have $A(u)$. Symbolically: $\exists x\,A(x) \Longrightarrow \exists x\big(\,A(x)\,\&\,\forall y \in x\,\neg A(y)\big)$

Note: we thus have the idea that for anything there is a first stage at which that thing happens.

We have two remaining problems:

(1) What are the properties of this hierarchy
(2) For how long does it go on?

Answer to (2) (part (a)).

1.6 Infinity. We need the natural numbers. So here they are :

$$\begin{aligned} 0 &= \emptyset \\ 1 &= \{0\} \\ 2 &= \{0,1\} \\ 3 &= \{0,1,2\} \\ &\vdots \end{aligned}$$

Note if we define $s(x) = x \cup \{x\}$ we have $s(n) = n + 1$ for each n.

We also want a set of natural numbers: $\mathbf{N} = \omega = \{0, 1, 2, \ldots\}$. Thus we would like to define

$$\omega = \bigcap\{x \mid 0 \in x \,\&\, \forall u \in x\; s(u) \in x\}.$$

But are there any infinite sets? Well, yes, – but we have to stipulate that they exist.

Infinity. $\exists x\, (\emptyset \in x \,\&\, \forall u \in x\; s(u) \in x)$.

(Note. V_ω is such an x.)

Other notions of infinity. (1) Infinite = not finite: "$\exists x\, \forall n \in \omega\; x \not\sim n$" wher $x \sim u$ if and only if $\exists f : x \longrightarrow y$ and f is a bijection. (In words: there is some x not in a bijection with any natural number.) But to say this we need functions – which we haven't yet defined.

(2) Dedekind Infinity: "$\exists x\, \exists u \subseteq x\, (x \neq u \,\&\, x \sim u)$." (In words: there is a set which has a proper subset of the same size as the set.)

If a set is Dedekind infinite it is infinite. To show, in the other direction that an infinite set is Dedekind infinite requires a small amount of the Axiom of Choice (to show that if x is not finite then x contains a countable subset).

1.7. The language of set theory and properties. Properties are the properties that are definable in a language, the language of set theory, LST, with sets as parameters. In fact we can choose a very simple language for the language of set theory.

(i) variables: u, v, x, z, ...
(ii) symbols for relations: $=$, $\in$
(iii) propositional logical connectives: & (and), v (or), $\longrightarrow$ (implies), $\neg$ (not), $\longleftrightarrow$ (if and only if)
(iv) quantifiers: $\forall x$ (for all x), $\exists x$ (there is some x).

Note: if we have a set a we have that $\{x \mid a \in x\}$ is a property and not a set.

With luck you understand this language informally and can argue using it. (For a more formal introduction to first order logic take a look at, *e.g.*, van Dalen's "Logic and Structure".)

We are also going to use some natural abbreviations:

$$
\begin{array}{ll}
x \neq z & \text{in place of } \neg(x = z) \\
x \notin z & \text{in place of } \neg(x \in z) \\
\forall x \in z\, \phi(x) & \text{in place of } \forall x\, (x \in z \longrightarrow \phi(x)) \\
\exists x \in z\, \phi(x) & \text{in place of } \exists x\, (x \in z\ \&\ \phi(x)) \\
\exists ! x\, \phi(x) & \text{in place of } \exists x\, (\phi(x))\ \&\ \forall z\, (\phi(z) \longrightarrow z = x)
\end{array}
$$

,

and so on and so forth. (Also I will often write "or" instead of " `v` " to avoid orthographic confusion between " `v` " and "v.")

Introduction of symbols for relations and functions. If I have a formula $\phi(x)$ in LST then I can introduce a new symbol $R_\phi(x)$ for a relation and an axiom $\forall x(R_\phi(x) \longleftrightarrow \phi(x))$. This doesn't change the expressive power of the language (what one can say in the language). For example: one can write $x \subseteq z$ if and only if $\forall u \in x\, u \in z$, where $\subseteq$ is our new symbol.

Similarly, if we have $\phi(x_0, \ldots, x_{k-1}, z)$ and we can show (from some axioms) that, writing $\overrightarrow{x}$ for $x_0, \ldots, x_{k-1}$, $\forall \overrightarrow{x}\, \exists ! z\, \phi(\overrightarrow{x}, z)$, we can introduce a symbol $f_\phi(z)$ for a function and a axiom $\forall \overrightarrow{x}, y\, (\phi(\overrightarrow{x}, y) \longleftrightarrow f(\overrightarrow{x}) = y)$. Again this doesn not change what the language can express.

I'm next going to give some examples of use of this language. As we have already encountered most of the Zermelo-Fraenkel axioms it seems like a good idea to use them as our first examples.

1.8 The axioms of Zermelo-Fraenkel (ZF) set theory

The axioms are (briefly):

Extension
Empty set
Pairs (if we have two sets we also have the set to which belong exactly the two sets)
Infinity
Unions
Power sets
Comprehension (or formation of subsets)
Collection (not discussed up to now)
$\in$-indution **or** Foundation: If A is non-empty then A has an element which is mininal in $\in$ ($\in$-minimal).

Zermelo produced all of the axioms except Collection and so the system without Collection is called Zermelo (Z) set theory.

[Bibliographic aside *re* history of development of axioms: see Michael Hallet, "Cantorian Set Theory and the Limitation of Size".]

Exercise. Show that Z holds in $V_{\omega+\omega}$.

Now I give them again, using LST. (I also include a couple of comments.)

Extension. $\forall x\, \forall y\, \big(\forall z\, (z \in x \longleftrightarrow z \in y\big) \longrightarrow x = y)$.

Empty set. $\exists x\, \forall y\, (y \notin x)$.

By the axiom of extentionality $\exists ! x\, \forall y\, (y \in x)$. Thus we can introduce a new symbol for a constant $\emptyset$ such that $\forall y\, y \notin \emptyset$. (A constant is merely a function of 0 variables.)

Pairs. $\forall x\, \forall y\, \exists! z\, \forall u\, (u \in z \longleftrightarrow u = x \vee u = y)$.

We can introduce a symbol for the function $\{.,.\}$ with $\forall u\, (u \in \{x, y\} \longleftrightarrow u = x \vee u = y)$. (The pair of two sets is unique by extension.)

Now we can see that $\{\emptyset\} \in \{\emptyset, \{\emptyset\}\}$ and so $1 \in 2$ (introducing the symbols 1 and 2 for constants in the obvious way for these sets).

Unions. $\forall x\, \exists z\, (u \in z \longleftrightarrow \exists y \in x\, (u \in y))$.

Again we can introduce a symbol, $\bigcup$ for a function and we have $\forall x\, (u \in \bigcup x \longleftrightarrow \exists y \in x\, u \in y)$.

Power Set. $\forall x\, \exists! z\, (y \in z \longleftrightarrow \forall u \in y\, (u \in x))$.

Once more we introduce a symbol $\mathcal{P}$ with the effect that $z \in \mathcal{P}(x) \longleftrightarrow z \subseteq x$.

Infinity. $\exists x\, (\emptyset \in x \;\&\; \forall y \in x\, (y \cup \{y\} \in x))$.

We introduce a symbol s for a new function and set $s(y) = y \cup \{y\}$, and define ω by

$$u \in \omega \longleftrightarrow \forall x\, ((\emptyset \in x \;\&\; \forall y \in x\, (y \cup \{y\} \in x)) \longrightarrow u \in x).$$

With two further steps we can say that $\omega = \bigcap\{x \mid \emptyset \in x \;\&\; \forall y \in x\, (y \cup \{y\} \in x)\}$.

Comprehension and classes. By a property of sets I mean some formula $\phi(\overrightarrow{x}, \overrightarrow{a})$ where $\overrightarrow{a} = a_0, \ldots, a_{n-1}$ are names for some specific sets. The property is true of $b_0, \ldots, b_{m-1} = \overrightarrow{b}$ if and only if $\phi(\overrightarrow{b}, \overrightarrow{a})$ is true. The *extension* of a property is the collection of sets of which the property is true. This collection is a subcollection of V, the collection of all sets. We call the extension of a property a *class*. Thus a class is something of the form $\{x \mid \phi(x, \overrightarrow{a})\}$.

Axiom of Comprehension. For all formulas $\phi(x, \overrightarrow{y})$ in LST we have

$$\forall x \forall a_0, \ldots, \forall a_{n-1}\, \exists y\, \forall z (z \in y \longleftrightarrow z \in x \;\&\; \phi(z, \overrightarrow{a})).$$

Note, this "axiom" is in fact an axiom schema: we have one axiom for each formula $\phi(x, \overrightarrow{y})$ in LST, and not merely one single axiom.

In effect this schema says that for every x we have that each $x \cap \{z \mid \phi(z, \overrightarrow{a})\}$ is a set. Thus a class $\{x \mid \phi(x, \overrightarrow{a})\}$ is a set (in the sense that $\exists z\, \forall x\, (x \in z \longleftrightarrow \phi(x, \overrightarrow{a}))$ if and only if $\{x \mid \phi(x, \overrightarrow{a})\} \subseteq V_\alpha$ for some α. *I.e.*, $\{x \mid \phi(x, \overrightarrow{a})\}$ is a set if and only if it is bounded in the hierarchy.

It is very useful (as we shall see) to be able to use various classes as informal abbreviations. Everyone does it and so shall we. However, you should bear in mind that the classes we use are only abbreviations and their use can always be eliminated in favour of talking about the formulars of LST that define them (*cf.* the above remarks about V).

For example, if $A = \{x \mid \phi(x, \overrightarrow{a})\}$ is a class, we can use $b \in A$ s an abbreviation for $\phi(b, \overrightarrow{a})$ being true, but one cannot "think about" whether $A \in A$ is true. It is not false, but *doesn't have any meaning*. Similarly, $P(A)$ does not have any meaning.

If you still have questions or doubts about how to use or think about classes wait a short while and after we have used them a little things should become clearer.

Foundation. For all formulas $\phi(x, \overrightarrow{a})$ in LST $\exists y\, \phi(y, \overrightarrow{a}) \longrightarrow \exists y\, \phi(y, \overrightarrow{a})\, \&\, \neg\exists z \in y\, \phi(z, \overrightarrow{a})$.

Using classes as abbreviations this says that if A is a class and is non-empty then A has an $\in$-minimal element.

How high is the hierarchy? (b) Suppose that $\forall x \in u\, \exists y\, \phi(x, y, \overrightarrow{a})$. For each x there is a first α_x such that $\exists y \in V_{\alpha_x}\, \phi(x, y, \overrightarrow{a})$ (by the axiom of foundation). So we have a function sending u to $\{\alpha_x \mid x \in u\}$ given by $x \mapsto \alpha_x$. Thus $\{\alpha_x \mid x \in u\}$ is a "set" of levels, and hence there is a β such that $\alpha_x < \beta$ for all $x \in u$. So we have $V_{\alpha_x} \in V_\beta$ for all $x \in u$. If we write v for V_β we see that $\exists v\, \forall x \in u\, \exists y \in v\, \phi(x, y, \overrightarrow{a})$.

But wait! We need an axiom (well, actually a schema of axioms, again) to do the "collecting" that gave us the set of levels we used above.

Axiom of Collection. For each formula $\phi(x, y, \overrightarrow{z})$ in LST (with free variables $\overrightarrow{z}$) we have the following:

$$\forall \overrightarrow{a}\, \forall u (\forall x \in u\, \exists y\, \phi(x, y, \overrightarrow{a}) \longrightarrow \exists v \forall x \in u\, \exists y \in v\, \phi(x, y, \overrightarrow{a})).$$

Surprisingly this schema implies that the hierarchy carries on for a very long time.

Note that feeding the "$\forall u$" through the "$\longrightarrow$" one gets that the axiom can be re-written as: for each formula $\phi(x, y, \overrightarrow{z})$ in LST and tuple $\overrightarrow{a}$ of sets one has

$$\forall x \in u\, \exists y\, \phi(x, y, \overrightarrow{a}) \longrightarrow \forall u\, \exists v\, \forall x \in u\, \exists y \in v\, \phi(x, y, \overrightarrow{a}).$$

Principle of Replacement For each formulas $\phi(x, y, \overrightarrow{z})$ in LST (with free variables $\overrightarrow{z}$) we have the following:

$$\forall \overrightarrow{a}\, (\forall u\, \exists! v\, \phi(u, v, \overrightarrow{a}) \longrightarrow \forall x\, \exists y\, \forall z\, (z \in y \longleftrightarrow \exists w \in x\, \phi(w, z, \overrightarrow{a}))).$$

Using classes as abbreviations again this says, "If F is a operation definable in V then for every set x we have que $F \restriction x$"x (the image of the restriction of F to x) is also a set."

Some authors prefer to use Replacement in place of Collection in formulating the axioms of ZF. I think that it is better to use Collection (I think it is simpler to understand what is going on in it, in part because one does not have to use classes to formulate it), but it does not make any difference formally which you use because Replacement follows from Collection and Comprehension.

(**Proof.** For each $\overrightarrow{a}$ the lhs gives for each x that we have $\forall u \in x\, \exists(!) v\, \phi(u, v, \overrightarrow{a})$. Using Collection we have $\exists s\, \forall u \in x\, \exists v \in s\, \phi(u, v, \overrightarrow{a})$. But the y that we need is a subcollection definable from s and so exists (is a set) by Comprehension.)

1.9. Other mathematially useful notions. Now we can introduce various other mathematically useful notions, giving definitions of them in LST.

Definition *Ordered pairs.* $(x, y) = \{\{x\}, \{x, y\}\}$.

Proposition. (a) For all x and y, (x, y) is a set.
(b) For all x, y, u and v, $(x, y) = (u, v) \longrightarrow x = u\ \&\ y = v$.

Proof. (a) By the axiom of Pairs (applied three times). (b) By extensionality.

Note: if $x = y$ the definition gives that $(x, x) = \{\{x\}\}$ and we have to check this case as well.

Definition. *Cartesian product of two sets.* $x \times y = \{(a, b) \mid a \in x \,\&\, b \in y\}$.

Proposition. For all x and y, $x \times y$ is a set.

Proof. For all x and y, $x \times y$ is a definable subcollection of $\mathcal{P}(\mathcal{P}(\bigcup\{x, y\}))$ and so is a set by the axioms of comprehension, power set, unions and pair.

Definition. *Relations and functions.* A relation on the product $x \times y$ is a subset of $x \times y$, *i.e.*, an element $\mathcal{P}(x \times y)$. So we can write, equally, $\mathrm{Rel}(x, y)$ $\mathcal{P}(x \times y)$ as an *aide memoire*. Functions are (just) certain relations. The collection of functions from x to y is written

$$^{x}y = \{f \in \mathrm{Rel}(x, y) \mid \forall a \in x\, \exists! b \in y\, (a, b) \in f\}.$$

Proposition. For all x and y, $\mathrm{Rel}(x, y)$ and ^{x}y are sets.

(We could carry on in this way, defining, for example, $\mathrm{dom}(z) = \{a \mid \exists b\, (a, b) \in z\} \subseteq \bigcup\bigcup z$ and $\mathrm{im}(z) = \{b \mid \exists a\, (a, b) \in z\} \subseteq \bigcup\bigcup z$, and afterwards proving that, given that $\mathrm{Rel}(z) \longleftrightarrow \forall x \in z$ $\exists a, b\, (x = (a, b))$, we have $\mathrm{Rel}(z) \longleftrightarrow z \subseteq \mathrm{dom}(z) \times \mathrm{im}(z)$.)

Let us write $f : x \longrightarrow y$ when $f \in {}^{x}y$, $\mathrm{dom}(f) = x$ and $\mathrm{im}(f) \subseteq y$. We would like to use the notation $f(a)$ when $f \in {}^{x}y$, $a \in x$ and $(a, b) \in f$. If we are adding a name for a function definable in the language we ought to give a meaning to any $f(a)$. So we set $f(a) = b$ if $a \in \mathrm{dom}(f)$, and $f(a) = \emptyset$ if $a \notin \mathrm{dom}(f)$. Here we are using f in two senses, one as the name of the set (the function from x to y), and the other as a new symbol in an extension of the language by a function definable in LST. So we need to set $f(a)$ (f in the second sense) equal to $\emptyset$ if $a \notin \mathrm{dom}(f)$ (f in the first sense). In practice we will rapidly come to ignore such distinctions.

Now we can define injective, surjective and bijective for functions $f \in {}^{x}y$, and so we can write the axiom of infinity, for example, as:

Infinity. $\exists x\, \forall n \in \omega \neg \exists f \in {}^{x}n\, f$ is bijective.

Now, also, we can introduce the Axiom of Choice.

Axiom of Choice. $\forall x\, ((\forall y \in x\, y \neq \emptyset) \longrightarrow \exists f \in {}^{x}\bigcup x\, \forall y \in x\, f(y) \in y)$.

This says that for all sets of non-empty sets there is a function that chooses a member of each one. We'll see more about AC in chapter 3.

Other similar useful definitions. If I and $\{y_i \mid i \in I\}$ are sets we have:

General Cartesian Products. $\Pi_{i \in I} y_i = \{f : I \longrightarrow \bigcup\{y_i \mid i \in I\} \mid \forall i \in I\, f(i) \in Y_i\}$.

Disjoint unions. $\Sigma_{i \in I} y_i = \{(i, a) \mid a \in y_i\}$. (Sometimes one sees $\dot{\bigcup}_{i \in I} y_i$ written for $\Sigma_{i \in I} y_i$.)

Quotient of a set by an equivalence relation. R is an equivalence relation on x if R is reflexive, $\forall y \in x\, (y, y) \in R$, R is symmetric, $\forall y, z \in x\, (y, z) \in R \longleftrightarrow (z, y) \in R$ and R is transitive, $\forall w, y, z \in x\, (w, y) \in R \,\&\, (y, z) \in R \longleftrightarrow (w, z) \in R$. If R is an equivalence relation on x we have $x/R = \{y \in \mathcal{P}(x) \mid \exists a \in x\, \forall b\, (b \in y \longleftrightarrow (a, b) \in R)\}$.

Proposition. Each $\Pi_{i \in I} y_i$, $\Sigma_{i \in I} y_i$ and x/R is a set.

Theorem. The collection of all sets, V, is closed under
(i) definable subsets
(ii) unions,
(iii) products indexed b sets,
(iv) disjoint unions indexed by sets,
(v) spaces of functions
(vi) power sets
(vii) quotients
and also contains $\mathbf{N}$ – *i.e.*, an object $\mathbf{N}$ with a function $s : \mathbf{N} \longrightarrow \mathbf{N}$ which satisfies Peano's axioms (which are give the rules that we expect for the conduct of the successor function on the natural numbers).

Using this theorem we can construct more or less any mathematical object (and it is easy to augment the theorem in order to show how to carry out other mathematical constructions in the theory of sets). These are no great shakes as you can see from the following examples.

We can define + and . on the natural numbers inductively. First of all $+ - n+(m+1) = s(n+m)$; and afterwards $. - n.(m+1) = n.m + n$. Then we can define $\mathbf{Z}$ as $\mathbf{N}\times\{0\}\dot{\cup}\{0\}\times\{1\}\dot{\cup}\mathbf{N}\times\{2\}$ ($\subseteq \mathbf{N}\times 3$), with $(i,y) < (j,z)$ if $i < j$, or $i = j = 0$ and $z < y$, or $i = j = 2$) and $y < z$, for example. (But clearly we could also choose any other reasonable scheme.) We can then give the laws of +, . on $\mathbf{Z}$. Next one can define $\mathbf{Q}$ as $\mathbf{Z}\times\mathbf{N}/\sim$ where $\sim$ is a relation $(a,b) \sim (c,d)$ if $a.d = b.c$. After this one can define $\mathbf{R}$ via Dedekind cuts. Or one could equally well, for example. define $\{r \in \mathbf{R} \mid 0 < r < 1\}$ via binary expansions, or via continued fractions, and take $\mathbf{R} = \mathbf{Z}\times\{r \in \mathbf{R} \mid 0 < r < 1\}$. (You can find more information about continued fractions in any good book on the elementary theory of numbers, such as, for example, Alan Baker's excellent *A Consise Introduction to the Theory of Numbers*, Cambridge University Press, pp. 41-42.)

As you will have seen by now, these constructions are more or less trivial and one has a great deal of freedom in the details of how one chooses to carry them out.