1.3 Set Theory I

As long as the terms of a mathematical theory are names of concrete objects – as concrete as mothers breast, the very first object that received a name in human languages - there is not much danger of formulating anything absurd. But on the other hand a language that deals only with these concrete things will not tell us much interesting stuff. In order to advance to more sophisticated utterances one has to admit new *mental objects*.

One may form a new mental object from given objects by grouping them. This new object is called a *set*. So for instance, if *a*, *b*, and *c* are objects the totality of these objects is a new object which is usually written putting the three names of the objects separated by commas in curly brackets: $\{a,b,c\}$. When one introduces this sort of new object a new type of statement is absolutely natural: the statement that a given object pertains to a set. So if M is a set and a an object, this object can be part of the set M. Such statement is written with the formula

$$a \in M \tag{1.3.1}$$

and in English one says "*a* is an element of *M*". The statement that *a* is not an element of M is written as

$$a \notin M \qquad (1.3.2).$$

The symbol "$\in$" has been introduced by Giuseppe Peano[1] to abbreviate the Greek word εστι (is).

When one introduces a new kind of object one has so say some words concerning the equality sign. It is not that our explanation of the symbol "=" has to be changed, but a discussion of the meaning of that symbol will actually be part of the definition of the new object. What do we actually mean by "the totality of the objects *a*, *b* and *c*"? Is the "totality of *a*, *b* and *c*" the same thing as the "totality of the objects *b*, *a* and *c*"? For sets one defines:

Two sets M and N are equal if and only if all elements of M are also elements of N and all elements of N are also elements of *M*.

We may write this definition as a formula:

$$\forall(\text{set } M)\forall(\text{set } N):\Big(\forall x:(x\in N \Leftrightarrow x\in M) \underset{Def.}{\Leftrightarrow} M=N\Big) \tag{1.3.3}$$

The formula (1.3.3) is the first axiom of set theory. It is called ***extensionality axiom*.** Note that we have modified our notation of logical quantifiers a little bit. In the quantification of the variables *M* and *N* we express that these variables are names of sets. The variable *x* may be any kind of object. It may also be a set.

Actually it would have been enough to write a simple implication in (1.3.3) because the other direction "$\Leftarrow$", which means the implication $N=M \Rightarrow \forall x:(x\in N \Leftrightarrow x\in M)$, is a consequence of the general rule (1.2.7).

An immediate consequence of this definition is that $\{a,b,c\}$ and $\{b,a,c\}$ are the same set. So the order of writing the elements does not matter.

---

[1] * 27 August 1858, †20 April 1932

It is convenient to define another type of statement that involves two sets. One says that a set N is a *subset* of a set M if and only if all elements of N are also elements of M. This statement is abbreviated with the sequence of symbols $N \subset M$.

$$N \subset M \quad \text{is an abbreviation of} \quad \forall x : \left(x \in N \Rightarrow x \in M \right) \tag{1.3.4}$$

Then one has $N = M$ if and only if $N \subset M$ and $M \subset N$.

The construction of a set by explicit presentation of objects is a rather primitive way of defining a set. It is much more elegant and interesting to specify the objects that pertain to a set by a description of their properties. This can be done with the help of statements. So one may define a set of all objects for which a certain statement is true. For instance, one may think of the set of all politicians that have been murdered. So Julius Caesar would be an element of this set. However, this way of defining a set contains certain dangers. The theory of sets had been invented by Georg Ferdinand Ludwig Philipp Cantor[2] in the 1870s in order to study certain infinite objects. Friedrich Ludwig Gottlob Frege[3] combined the theory with mathematical logic. The theory was accepted by very important mathematicians (for instance Julius Wilhelm Richard Dedekind[4] and David Hilbert[5]) and violently rejected by others (Leopold Kronecker[6] and Jules Henri Poincaré[7]). These controversies depressed Cantor considerably. An argument raised by Bertrand Arthur William Russell[8], which is now known as Russell's Paradox, raised even more doubts concerning the validity of set theory. Today set theory can be formulated in a sound and safe way and it is one of the cornerstones of modern mathematics. Russell's Paradox concerns exactly the definition of sets by means of statements. It goes as follows. Let $R_{Russell}$ be the set whose elements are all sets that are not their one element. So $R_{Russell}$ consists of all sets x such that the statement $x \notin x$ is true. Now consider the following question: is $R_{Russell} \in R_{Russell}$? If we assume $R_{Russell} \in R_{Russell}$ then, by definition of $R_{Russell}$, it should be a set that is not its own element, which contradicts the assumption $R_{Russell} \in R_{Russell}$. If, on the contrary, we assume $R_{Russell} \notin R_{Russell}$ then, by definition of $R_{Russell}$ it should be an element of $R_{Russell}$, which contradicts the assumption $R_{Russell} \notin R_{Russell}$. So whatever we assume as true, we end up with a contradiction.

This situation seems to show that the idea of sets is flawed. But a closer look at the original idea of set shows that Russell's argument is not valid. We started the second paragraph of this section with the sentence: "One may form a new mental object from <u>given</u> objects by grouping them." The objects that are grouped to form a set have to be existing objects. Once a set is defined this is a new existing object. But the formation of a set is always based on previously existing objects. Therefore sets have to be thought of as organized in a hierarchy. This hierarchy starts with some concrete objects, which are not sets. From these, which were called "urelemente", here we shall call them *basic elements*, one may safely form any kind of sets. These sets are new mental objects,

---

[2]* March 3 [O.S. February 19] 1845. †January 6, 1918
[3] *8 November 1848, †26 July 1925
[4] *October 6, 1831, † February 12, 1916
[5] January 23, 1862 †February 14, 1943
[6] *December 7, 1823, † December 29, 1891
[7] *29 April 1854, † 17 July 1912
[8] *18 May 1872, † 2 February 1970

which together with the basic elements can again be grouped to form sets of a higher sphere, which again may be grouped to sets of an ever higher sphere and so on. Ernst Friedrich Ferdinand Zermelo[9] and Adolf Abraham Halevi Fraenkel[10] formulated axiomatic rules that permit formation of sets without the danger of building any monsters like $R_{Russell}$.

Before we explain these axioms a note on notation is appropriate. Formulas of the type $a \in b$ or $a \notin b$ were declared to be syntactically correct only in the case where b is a set. This may cause inconvenient situations. We shall extend the notation to the case where b is not a set. If b is not a set we define $a \in b$ to be false for all a and $a \notin b$ to be true for all a.

First of all, the initial idea that one may join given things to form a set is expressed with the ***pairing axiom*.** For any objects *a* and *b* there exists a set that contains exactly *a* and *b* as elements.

$$\forall a \forall b \exists (\text{set}\, M)\, \forall x : \left(x \in M \Leftrightarrow x = a \vee x = b \right) \qquad (1.3.5)$$

With the extensionality this set is uniquely determined by *a* and *b*. It is written as $\{a, b\}$. Note, that one may also have $a = b$. So for any a one also has a set $\{a\}$, which is uniquely determined by a. The case of more than two objects will be indirectly taken care of with the remaining axioms.

The formation of sets by means of a statement is restricted to subsets and is expressed by the ***Aussonderungsaxiom***[11] or ***axiom of restricted comprehension*** or ***subset axiom.*** It tells us that for any set M and any formula Φ that contains the free variables *M, a, b, …, x* but that does not contain the variable S freely there is a set S (which depends on *a, b, …*) whose elements are also elements of M and that satisfy the formula Φ.

$$\forall (\text{set}\, M)\, \forall a \forall b \forall c .....\exists (\text{set}\, S)\, \forall x : \left(x \in S \Leftrightarrow \left(x \in M \wedge \Phi \right) \right) \qquad (1.3.6)$$

Extensionality implies that this set is uniquely determined by *M* , *a, b, …* and by the formula Φ. Usually this set is written as $\{x \in M \mid \Phi\}$. One special case, which at first sight seems meaningless, but which determines a frequently used set, is obtained with the formula $x \neq x$. This gives a set $\{x \in M \mid x \neq x\}$ that contains no elements. By extensionality this empty set is unique. It is usually abbreviated with the symbol $\varnothing$. Another important application of the subset axiom is the definition of the relative complement. If *M* and *N* are sets one defines the complement of *N* relative to *M* as

$$M \setminus N \underset{Def.}{=} \{x \in M \mid x \notin N\} \qquad (1.3.7)$$

[9] *1871, †1953

[10] *February 17, 1891, † October 15, 1965

[11] From German aussondern = to separate out

The ***union axiom*** tells us that for any set M there is a set $\bigcup M$ whose elements are the elements of elements of *M*.

$$\forall(\text{set}\, M)\exists(\text{set}\, N)\forall x:\big(x\in N \Leftrightarrow (\exists a:(a\in M \wedge x\in a))\big) \qquad (1.3.8)$$

This set is unique and it is usually written as $\bigcup M$ and called the union of *M*. Especially if M has two elements each one being a set $M=\{A,B\}$ (A and B sets) the union $\bigcup M$ consists of all objects that are elements of *A* or *B*. This set is usually written as $A\cup B$:

$$A\cup B \underset{Def.}{=} \{x \mid x\in A \vee x\in B\} \qquad (1.3.9)$$

If M is a set whose elements are all basic elements, that is no element is a set, the union of *M* is empty: $\bigcup M=\varnothing$. The union axiom together with the pairing axiom permits one to aggregate more than two elements in an explicit construction of a set. So for instance, to build the set $\{a,b,c\}$ one first uses the pairing axiom to build the sets $\{a\}$ and $\{b,c\}$, then one uses the pairing axiom again to form $\{\{a\},\{b,c\}\}$ and finally the union of this set is the desired $\{a,b,c\}$.
The existence of the union had to be formulated as an axiom. When one replaces the "$\vee$" by an "$\wedge$" in the formula (1.3.9) one gets a set whose existence can be proven. This set is called the intersection of the sets A and *B* and it is written as $A\cap B$:

$$A\cap B \underset{Def.}{=} \{x\in A\cup B \mid x\in A\wedge x\in B\} \qquad (1.3.10)$$

The existence is a consequence of the subset axiom. One can also define the intersection of the elements of a set:

$$\bigcap M \underset{Def.}{=} \{x\in \bigcup M \mid \forall a:(a\in M \Rightarrow x\in a)\} \qquad (1.3.11)$$

If the set M contains basic elements (urelemente) then $\bigcap M$ is empty.

From an object a we can form a set $\{a\}$. But is this new object really different form *a*? In the case that $a=\varnothing$ we can actually show that it is different. The set $\{\varnothing\}$ has one element, which is the empty set. Therefore $\{\varnothing\}$ cannot be the empty set, which by definition does not have any element. So $\varnothing\neq\{\varnothing\}$. But for the general case we have no argument that shows $a\neq\{a\}$. This statement can be shown to be true with the help of the ***axiom of foundation*** or ***regularity axiom*.** The idea behind this axiom is the following: We may find ever higher and higher levels of hierarchy among the elements of a set. So for instance we may construct sets that contain the following kind of elements: $\{\varnothing,\{\varnothing\},\{\{\varnothing\}\},\{\{\{\varnothing\}\}\},............\}$. But the idea that all construction of sets starts at some fixed stock of concrete objects implies that a going back to lower and lower levels of hierarchy must eventually stop at some point. This motivates to require that there exists an element in any non-empty set M that does not contain elements of *M*. Written as a formula this axiom reads as follows:

$$\forall(\text{set } M):\left(M \neq \varnothing \Rightarrow \exists x:\left(x \in M \wedge \forall y:(y \in M \Rightarrow y \notin x)\right)\right) \qquad (1.3.12)$$

With this axiom and with the pairing axiom we can show the following theorem:

Theorem1.3.1: $\forall(\text{set } N): N \notin N$

Proof. Let a set N be given arbitrarily. With the pairing axiom we can form the set $M=\{N\}$. *M* is not empty. Then, according to the regularity axiom, this set *M* must contain an element x such that for all elements y in M one has $y \notin x$. But there is only the element N in *M* so it follows $N \notin N$.

An immediate consequence of that theorem is

Theorem 1.3.2: $\forall x: x \neq \{x\}$

Proof. Let *x* be given arbitrarily. Applying theorem 1.3.1 to the set $\{x\}$ one sees that $\{x\} \in \{x\}$ is false. On the other hand, $x \in \{x\}$ is true. The first formula can be obtained from the second one by substituting $\{x\}$ in the place of the expression *x* on the left hand side of the symbol "$\in$". Then the rule (1.2.6) gives $x \neq \{x\}$.

The proof of the following theorem is left as an exercise:

Theorem 1.3.3: $\forall(\text{set } M)\forall x:(x \in M \Rightarrow M \notin x)$

Next we have the ***power set axiom.*** This simply states that the totality of all subsets of a set is also a set:

$$\forall(\text{set } M)\exists(\text{set } N)\forall x:(x \in N \Leftrightarrow x \subset M) \qquad (1.3.13)$$

This set is called the powerset of M and it is written as $\mathrm{P}(M)$.

Further one requires that one obtains a set if one replaces the elements of a given set M by other objects that are uniquely determined by the elements of the given set. This determination of new objects is established by means of a formula $\Phi[x,y]$ that contains the free variables x and y , where "*x*" is a name of elements of M. For every $x \in M$ there should exist a unique y such that $\Phi[x,y]$ is true. In order to write this axiom in a comprehensible manner it is convenient to introduce an abbreviation for the existence of a unique *y* such that $\Phi[x,y]$ is true:

$$\exists! y:\Phi[x,y] \quad \underset{Def.}{\Leftrightarrow} \quad (\exists y:\Phi[x,y]) \wedge \left(\forall a \forall b:(\Phi[x,a] \wedge \Phi[x,b]) \Rightarrow a=b\right)$$

(1.3.14)

Further we shall define two more abbreviations:

$$\forall(a \in M):\Phi \quad \underset{Def.}{\Leftrightarrow} \quad \forall a:(a \in M \Rightarrow \Phi) \qquad (1.3.15)$$

$$\exists(a \in M):\Phi \quad \underset{Def.}{\Leftrightarrow} \quad \exists a:(a \in M \wedge \Phi) \qquad (1.3.16)$$

Replacement axiom: For any formula $\Phi[x,y]$ in which *x*, *y*, *a*, *b*, …and *M* appear freely but *N* does not appear freely the following formula is true:

$$\forall a \forall b .. \forall (\text{set } M) : \left(\left(\forall (x \in M) \exists ! y : \Phi[x, y] \right) \Rightarrow \exists (\text{set } N) \forall y : \left(y \in N \Leftrightarrow \exists (x \in M) : \Phi[x, y] \right) \right) \tag{1.3.17}$$

A simple application of the replacement axiom and the union axiom is the definition of the union of a family of sets. Let I be a non-empty set and suppose that one has a unique set A_i for every $i \in I$. Such a collection of sets is called a *family of sets* and the set I is called the *index set of the family*. Then, the replacement axiom tells us that there exists a set M whose elements are exactly the sets A_i.

$$M = \{A_i \mid i \in I\} \tag{1.3.18}$$

The union of the family is then defined as the union of M:

$$\bigcup_{i \in I} A_i \underset{Def.}{=} \bigcup \{A_i \mid i \in I\} \tag{1.3.19}$$