

Elements of Set Theory

Péter Mekis
Department of Logic, ELTE Budapest

February 24, 2015

Preface

This is an as yet incomplete draft of what is going to be the lecture notes for a course under the same title for students of the *Logic and the Theory of science* MA program. The text is written for students that are logically and philosophically minded, but do not necessarily have a mathematical background. However, it requires familiarity with the basics of first-order logic. A reader who is skilled in reading mathematical texts will find most of the proofs too detailed. Note that any proof can be skipped and taken as an exercise to avoid boredom.

Instead of the more well-known system of Zermelo–Fraenkel set theory, these notes use the framework of Gödel–Bernays set theory (GB). Beyond the personal taste of the author, the reason for this choice is that there is a follow-up of this course under the title *Alternative set theories*, the basic task of which is to investigate the philosophical motivations and implications of various versions of the set/class distinction, which is at the heart of any axiomatic set theory, and which is more convenient to discuss in a language where one can quantify over classes.

A mathematically minded reader may be surprised by the length of discussing certain basic topics, even without any apparent practical profit, and even at the cost of giving a rather sketchy presentation of more advanced results. An example is the detailed comparison of three different ways to represent natural numbers in GB. The reason is that the discussion aims at foundational questions, and how to represent numbers by sets adequately is one of the great questions of the foundational studies. Also, note that different representations of numbers fit different set theoretical frameworks, so numbers are nice and relatively simple objects for a comparative study of alternative set theories. In other words: our aim is establishing theories, not establishing results within theories.

The author apologizes for all mistakes due to his poor English, and also for the innumerable number of typos due to the work being under preparation. He hopes that the number of major mistakes is decreasing in the process of revisions. Questions, corrections and suggestions are welcome to mekis.peter@gmail.com.

1 Introduction: From Hilbert’s Hotel to Cantor’s Paradise

As William Weiss put it, “set theory is the true study of infinity.” Introducing a technical concept of infinity, free from any theological or metaphysical burden, is a major achievement of human culture, made possible by a systematic study of sets (classes, collections). In this introductory section we consider some basic arguments in this area in an intuitive manner. In the subsequent sections we will develop a technical framework in which these intuitive results can be reconstructed with more precision.

A first and seemingly paradoxical property of an infinite collection is that it's in some sense resistant to adding or withdrawing elements. This was already seen by Galileo, and systematically discussed in Bernard Bolzano's *Paradoxes of Infinity*. We present it in its popular form, as a thought experiment about an infinite hotel, the hotel's many guests, and its smart manager.

In Hilbert's Hotel, named after the German mathematician David Hilbert, the rooms are numbered in the usual way, except that there is no largest room number. Right after room number 100, we find room number 101; right after room 1000, there is room 1001; and so on, *ad infinitum*. Each one is a single room, suitable to accommodate one guest only. At the beginning of our story, the hotel is full; each room is taken by one guest.

On the first day, a new guest arrives and asks the hotel manager for a room. The receptionist solves the problem quickly. She asks each guest to move to the subsequent room; from room 1 to room 2, from room 2 to room 3, and so on, from room n to room $n + 1$. The new guest takes room 1. No one had to move out, one person moved in, and there is still one and only one person in each room.

On the second day, an infinite bus of new guests arrives; it has no largest seat number, and all the seats are taken. Can the new guests fit in the hotel? Yes; the receptionist asks each guest of the hotel to move to the room the number of which is twice the number of her original room. Thus the guest in room 1 moves to room 2, the guest in room 2 moves to room 4, the guest in room 3 moves to room 6, and so on, the guest in room n moves to room $2n$. Now all the even numbered rooms are taken, and all the odd numbered rooms are free, so the new guests can move in them; from seat 1 to room 1, from seat 2 to room 3, from seat 3 to room 5, and so on, from seat n to room $2n - 1$.

On the third day, an infinite sequence of infinite buses arrive to the hotel; there is a bus number one, a bus number two, and so on; there's no bus with a largest number. Can they all fit in the hotel? After a bit of thinking the receptionist creates a diagonal ordering of all the guests as shown in the following tabular. (The number after the arrow in the lower index is the number of the destination room.)

hotel	1 \rightarrow 1	2 \rightarrow 3	3 \rightarrow 6	4 \rightarrow 10	5 \rightarrow 15	6 \rightarrow 21	7 \rightarrow 28	...
bus 1	1 \rightarrow 2	2 \rightarrow 5	3 \rightarrow 9	4 \rightarrow 14	5 \rightarrow 20	6 \rightarrow 27	7 \rightarrow 35	...
bus 2	1 \rightarrow 4	2 \rightarrow 8	3 \rightarrow 13	4 \rightarrow 19	5 \rightarrow 26	6 \rightarrow 34	7 \rightarrow 43	...
bus 3	1 \rightarrow 7	2 \rightarrow 12	3 \rightarrow 18	4 \rightarrow 25	5 \rightarrow 33	6 \rightarrow 42	7 \rightarrow 52	...
bus 4	1 \rightarrow 11	2 \rightarrow 17	3 \rightarrow 24	4 \rightarrow 32	5 \rightarrow 41	6 \rightarrow 51	7 \rightarrow 62	...
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

The above thought experiment suggests a conclusion that in the case of an infinite collection of items (guests, rooms, seats, etc.) expansion (e. g. by adding new guests) or reduction (e. g. by emptying even numbered rooms) does not affect the collection's size; any infinite collection has the same number of elements as any other. We will see shortly strong arguments against this position; but before that, let us distill preliminary definitions of some basic concepts from the above discussion.

The thought experiment also suggests two different definitions of infinity. In one sense, *a set is infinite if and only if it has at least as many elements as the set of the natural numbers*. This is behind the idea that there is no largest room number in the hotel. If there was a largest room number, it would also be the number of the rooms, thus making the hotel finite. Since there is no largest room number, all the natural numbers are used in enumerating them; thus there are just as many rooms in the hotel as natural numbers. The other suggested definition exploits the idea that the hotel manager used in accommodating new guests in a full hotel; *a set is infinite if and only if it is possible to extract elements from it so that there are as many elements remaining as there were originally*.

In set theory, the first version became the standard definition of infinity; but the latter, labelled as Dedekind infinity (after the German mathematician Richard Dedekind) is also common. We are not yet in the position to tell whether they introduce the same concept or different ones; or, to phrase the question a bit more precisely, under what conditions do

the two concepts fall together. (Later we will see that in the standard axiomatic framework of set theory this depends on whether we assume a weak version of the axiom of choice, called the principle of countable choice.) But we can observe that both definitions rely on the relational concept *as many as*; a proper definition of which will reflect the very essence of the thought experiment.

Each trick of the hotel manager established one-one correspondences between sets. Based on this strategy, we can say that *a set has as many elements as another one if and only if there is a one-one correspondence between them*. This idea is sometimes called Hume's principle, and although it is not entirely clear whether Hume actually held it regarding sets (he applied it somewhat confusingly to numbers), it was present in Gottlob Frege's *The Foundations of Arithmetic*, and several other early investigations on the numerosity of sets. When a set has as many members as another one, we call them equinumerous. Intuitively, equinumerosity is an equivalence relation; it is reflexive, symmetrical and transitive. This will be the case in the standard axiomatization of set theory, too, as well as in many alternative axiomatic frameworks.

From a set-theoretical point of view, a one-one correspondence from a set H to the set G may be regarded as a set, too; that is, a set f of ordered pairs that satisfy the following conditions:

1. the first members of each pair in f are from H ;
2. each element of H occurs as a first member in exactly one pair in f ;
3. the second members of each pair in f are from G ;
4. each element of G occurs as a second member in exactly one pair in f .

As we have seen, the Hilbert's Hotel thought experiment suggests that any infinite set has as many members as any other. Let us now consider a counterargument against this, in the form of an additional episode to the Hotel saga. On the fourth day, after all the guests from the previous day left the hotel, a special infinite bus arrives, full of passengers. The seat numbers in this bus are real numbers between 0 and 1 (the latter reserved for the driver). There is a seat numbered as $2/3$ and one numbered as $4/5$; and in between them, there are seats number $\sqrt{2}/2$, $3/4$, $\pi/4$, etc. Some of these numbers, like $3/4$, are fractions of two integers; they are called rational numbers. Others, like $\sqrt{2}/2$, are roots of polynomial equations with integer coefficients (thus $\sqrt{2}/2$ solves the equation $2x^2 - 1 = 0$); these are the algebraic numbers, including the rationals. Yet others, like $\pi/4$, lack this connection with polynomials $\pi/4$ (as it was proved in the case of $\pi/4$ by Ferdinand von Lindemann); they are called transcendental.

We take it for granted that all of these seat numbers are expressible by infinite sequences of digits; $3/4$ as $0.75000\dots$, $\sqrt{2}/2$ as $0.70710\dots$, and $\pi/4$ as $0.78539\dots$ (in the first case the three dots represent an infinite sequence of zeros; in the other two cases, various other digits without periodicity). We will identify seat numbers with their decimal expansions. To avoid notational ambiguity, we don't allow for infinite sequences of nines in the expansions, since, for example, $0.74999\dots$ (where the dots stand for further nines) denotes the very same number as $0.75000\dots$ above.

We will prove that the passengers of the bus cannot be accommodated in the hotel. We do this by means of an indirect argument; that is, we assume (towards a contradiction) that the whole busload of passengers has been successfully accommodated in the hotel. The room arrangement will be something like the following, with the room numbers on the left, and the seat numbers on the right:

$$\begin{array}{l}
1 \longrightarrow 0. 7 5 0 0 0 \dots \\
2 \longrightarrow 0. 7 0 7 1 0 \dots \\
3 \longrightarrow 0. 7 8 5 3 9 \dots \\
4 \longrightarrow 0. 3 9 2 2 8 \dots \\
5 \longrightarrow 0. 3 9 2 2 3 \dots \\
\vdots \qquad \qquad \qquad \vdots \qquad \qquad \qquad \ddots
\end{array}$$

By hypothesis all seat numbers occur in the list. We construct one that isn't there. The procedure runs as follows: take the n th digit (after the decimal point) of the n th member of the list, and if it is less than 5, substitute 5 for it; otherwise substitute 4 for it. This procedure gives us all the digits of a new number c as follows:

$$\begin{array}{l}
1 \longrightarrow 0. \mathbf{7} 5 0 0 0 \dots \\
2 \longrightarrow 0. 7 \mathbf{0} 7 1 0 \dots \\
3 \longrightarrow 0. 7 8 \mathbf{5} 3 9 \dots \\
4 \longrightarrow 0. 3 9 2 \mathbf{2} 8 \dots \\
5 \longrightarrow 0. 3 9 2 2 \mathbf{3} \dots \\
\vdots \qquad \qquad \qquad \vdots \qquad \qquad \qquad \ddots \\
c = \quad 0. 4 5 4 5 5 \dots
\end{array}$$

By hypothesis c must be in the list; but it isn't, because it differs from each member on the list in at least one digit. We arrived at a contradiction, so our initial assumption must be rejected.

On the other hand, it is easy to see that the set of the passengers contains a part that is equinumerous with the set of the hotel rooms. For example, take the passengers with seat numbers of the form $1/n$ for positive integers n ; moving the passenger on seat $1/n$ to room n amounts to a one-one correspondence between the two sets.

We call such parts of a set—which themselves are sets, too—subsets. We can phrase a definition of the term as follows: *a set G is a subset of a set H if and only if every element of G is also an element of H .* In this sense every set is a subset of itself; and the empty set is a subset of every set. (The latter is a very useful construction included into the universe of sets in virtually all set theoretic frameworks; it is, e. g., the set of all natural numbers that are neither even nor odd.)

How many subsets does a set have? The question has a simple answer in the realm of finite sets. As an example, consider a set with three members, $\{a, b, c\}$. Its subsets include the singleton sets $\{a\}$, $\{b\}$, $\{c\}$, the two-membered sets $\{a, b\}$, $\{b, c\}$, and $\{a, c\}$, the set $\{a, b, c\}$ itself, and the empty set; eight subsets altogether. Or, counting more systematically, with each of the three members of $\{a, b, c\}$ there come two possibilities; it is either included or not included as an element of a particular subset. This amounts to $2 \times 2 \times 2 = 2^3 = 8$ possibilities altogether. Similarly, in the case of an n -membered set, there are 2^n subsets. Now for any positive number n , 2^n is larger than n ; therefore a finite set has always more subsets than elements. This can be seen by simple arithmetic induction; $1 < 2^1$, and for every n , if $n < 2^n$, then

$$n + 1 \leq n + n < 2^n + 2^n = 2^{n+1}.$$

How about infinite sets? Do they have more subsets than elements? Let us call the set of all subsets of a set its power set. We will prove that (a) a set cannot be equinumerous with its power set; and (b) the power set of any set has a subset that is equinumerous with the original set. These two clauses are known as *Cantor's theorem*. Let's start with (b), and consider the set G of all singleton subsets of a set H . Obviously there's a one-one correspondence between the two sets; the one that assigns to each member a of H its singleton $\{a\}$.

As for (b), our proof will take the form of an indirect argument. Suppose (towards a contradiction) that for some set H and its power set G , there is a one-one correspondence between H and G . For each member a of H , its pair A is a subset of H . Thus for each a

in H here are two possibilities: (1) a is in A ; (2) a is not in A . By these possibilities we can divide H into two halves. Let C be the set of those elements a of H which are not elements of the corresponding subset A .

Now C is a subset of H , too, and thus it is an element of its power set G . Since by assumption we have a one-one correspondence between H and G , C must be assigned to some element c in H . Again, c either is or isn't an element of C ; one of the above possibilities must hold. If (1) is the case, then by definition c is in C , so by the definition of C c is not in C , which is a contradiction. On the other hand, if (2) is the case, then c is not in C , which means that c satisfies the defining condition of C , therefore c is in C , which is, once again, a contradiction. There are no more possibilities; so the indirect assumption must be rejected.

Let us call a set G larger than a set H if the H is not equinumerous with G , but H is equinumerous with a subset of G . In these terms Cantor's theorem implies that there is no largest set; for every set H , finite or infinite, there is a set G that is larger than it. Can we conclude from this that in terms of largeness, there is an infinite hierarchy of infinite sets, the same way as there is an infinite hierarchy of finite sets? Not yet. First we have to prove that two sets cannot be mutually larger than each other. This result is widely known as the Schröder–Bernstein theorem. The proof we present was given by the Hungarian mathematician Gyula König.

Let H and G be arbitrary sets, so that there is a subset H' of H and a subset G' of G such that H is equinumerous with G' and G is equinumerous with H' . Based on these conditions we prove that H is equinumerous with G . Let f be a one-to-one correspondence between H and G' ; and let g be a one-to-one correspondence between G and H' . Our strategy is to partition H and G into smaller sets between which either f or g work as one-one correspondences; and to construct a one-one correspondence e from these partial correspondences. The definition of the parts will be simpler if we introduce another concept: the image of a set H with regards to a correspondence f is the set of all elements that occur as the second member of a pair in f , the first member of which is from H . In other words, the image of H with regards to f is the set of the values assigned to elements of H by f . In this sense G' is the image of H with regards to f ; and H' is the image of G with regards to g .

Consider now the following sequences of sets. Let H_1 be H , and let G_1 be G ; and for every positive integer n , let H_{n+1} be the image of G_n with regards to g ; and let G_{n+1} be the image of H_n with regards to f . This simultaneous definition makes, of course, H_2 to be the same as H' , and G_2 to be the same as G' ; and generates an infinite parallel sequence of sets in both sides in such a way that for every positive integer n , H_{n+1} is a subset of H_n , and G_{n+1} is a subset of G_n .

Now we can define the partitions of both sets as follows: for every positive integer n , let A_n be the set of those elements of H_n which are not in H_{n+1} ; and let B_n be the set of those elements of G_n which are not in G_{n+1} . It may or may not be the case that every element of H is in A_n for some n , and every element of G is in B_n for some n . So let A^* be the (possibly empty) set of those elements of H that are not in A_n for any n ; and let B^* be the (possibly empty) set of those elements of G that are not in B_n for any n .

Now we can start spotting equinumerous pairs in these sequences. By definition, an even-numbered member of the A -series is the image of the preceding member of the B -series with respect to g ; in other words, the restriction of g to B_{2n-1} is a one-one correspondence between B_{2n-1} and A_{2n} . The very same way, an even-numbered member of the B -series is the image of the preceding member of the A -series with respect to f ; in other words, the restriction of f to A_{2n-1} is a one-one correspondence between A_{2n-1} and B_{2n} . And, also as a consequence of their definitions, the restriction of f to A^* is a one-one correspondence between A^* and B^* (and *vice versa*, the restriction of g to B^* is a one-one correspondence between B^* and A^*).

Based on the above considerations, we can define a one-one correspondence e between H and G as follows. If an element a of H is in A_{2n} for some n , then let e assign to a

the element b of G to which g assigns a . If a is in A_{2n+1} for some n or a is in A^* , then let e assign to a the same element b that f assigns to it. thus defined, e is a one-one correspondence between H and G , demonstrating the equinumerosity of H and G , and thereby completing our proof of the Schröder–Bernstein theorem.

Now we can safely conclude from the above considerations that there is an infinite hierarchy of infinite sets. This hierarchy is often referred to as Cantor’s paradise.

We have to take the results of this section with some care. In what sense have we proved the existence of a set larger than the set of natural numbers, or the existence of an infinite hierarchy of infinities? We certainly didn’t, in the sense that anyone who reads these arguments should be convinced of the existence of this hierarchy. The arguments were based on intuitive assumptions, often left implicit. Let us enumerate some of the assumptions necessary to conclude that Cantor’s theorem is true:

- (1) there are sets of sets, not only sets of things that aren’t sets;
- (2) the subsets of any set form a set;
- (3) you can specify a subset of a set by means of any well-defined property.

Of these assumptions, (1) is accepted in virtually all axiomatizations of set theory; assuming the iterability of the elementhood relation is part of the reasons why we call a theory a set theory. But, of course, it is by no means necessary to discuss sets in a set-theoretical framework. (One can choose, for example, monadic second-order logic as a framework to discuss sets, where the values of predicate variables are sets of individuals, but there’s no way to talk about sets of sets.) As for (2), it is also assumed in most axiom sets of set theory, but not everywhere. (There are small theories with infinite sets that have no power sets.) Finally, while (3) is a core assumption in standard set theory and its akin, it’s rejected in some alternative set theories like New Foundations and related systems, or the various versions of positive set theory favor of other principles.

Believing in the existence of Cantor’s paradise, whatever the terms “believing” and “existence” may mean in this context, amounts to a metaphysical commitment. Technical considerations cannot solve metaphysical problems. The best that we can hope from a set-theoretical discussion of sets is making as clear as possible the consequences of such commitments, thus giving a technical basis for one’s choice of metaphysical assumptions. As an extreme case, by showing that a certain set of assumptions leads to contradiction is in this sense a major achievement. This is what we will do in the next section regarding a very appealing system of set-theoretical axioms.