Department of Philosophy - Utrecht University

Zermelo and the Skolem Paradox

D. van Dalen

θ

π

Logic Group
Preprint Series
No. 183
February 1998

Utrecht Research
Institute for
Philosophy

ISBN 90-393-1725-9
ISSN 0929-0710
Dr. A. Visser, Editor

Zermelo and the Skolem Paradox

Dirk van Dalen[1] and Heinz-Dieter Ebbinghaus[2]

Dedicated to Mrs. Gertrud Zermelo on the occasion of her 95th birthday

On October 4, 1937 Zermelo wrote down a hitherto unpublished manuscript entitled "Der Relativismus in der Mengenlehre und der sogenannte Skolemsche Satz" ("Relativism in Set Theory and the so-called Theorem of Skolem")[3] in which he gives a refutation of "Skolem's paradox", i.e., the fact that Zermelo-Fraenkel set theory – guaranteeing the existence of uncountably many sets – has a countable model. Compared with what he wished to disprove, the argument fails. However, at a second glance, it strongly documents his view of mathematics as based on a world of intuitively given objects that could only be grasped adequately by infinitary means.

Whereas the Skolem paradox was to raise a lot of concern in the twenties and the early thirties, it seemed to be settled when Zermelo wrote his paper, namely in favour of Skolem's approach, thus also accepting the noncategoricity and incompleteness of the resulting axiom systems. So the paper might be considered a late-comer in a community of logicians and set theorists who mainly followed finitary conceptions, in particular emphasizing the role of first-order logic (cf. [Moore 1980]). However, Zermelo never shared this viewpoint: In his first letter to Gödel from September 21, 1931 (cf. [Dawson 1985]) he had written that the Skolem paradox rested on the erroneous assumption that every mathematically definable notion should be expressible by a finite combination of signs, whereas a reasonable metamathematics would only be possible after this "finitistic prejudice" would have been overcome, "a task I have made my particular duty".

Our paper is organized as follows. In a first section we trace the discussion that originated in the discovery of the Skolem paradox. We then, in Section 2, describe Zermelo's view concerning the nature of mathematics. A part of the citations is taken from still unpublished documents of the Nachlass. Having thus provided for the necessary background, in Section 3 we give an annotated translation of the 1937 manuscript. An appendix contains the original German version.

[1]University of Utrecht, Dept. of Philosophy, P.O. Box 80.126, 3508 TC Utrecht, The Netherlands.

[2]University of Freiburg, Inst. of Mathematical Logic, Eckerstrasse 1, 79104 Freiburg, Germany.

[3]The three page handwritten paper is contained in the Zermelo Nachlass in the main library of the Universität Freiburg. We thank the Universitätsbibliothek for allowing us to publish it here and to cite from other papers in the Nachlass. Warm thanks go to Mrs. Gertrud Zermelo for kindly supporting work on her late husband.

1. The Skolem Paradox

The paper "Einige Bemerkungen zur axiomatischen Begründung der Mengenlehre" of Skolem [Skolem 1923] contains a number of important observations which have considerably influenced the development of set theory, and also that of the foundations of mathematics.

In a way the paper is the starting point of post-Zermelo axiomatic set theory. The particular issue is the "relativity of set theoretic notions, which is unavoidable in each consistent ("konsequent") axiomatization". The phenomenon that Skolem refers to has become known as "Skolem's paradox". The paradox almost immediately became the subject of discussions in foundational circles, where it caused a good deal of confusion.

Skolem handled the phenomenon in a precise and definitive manner. He gave a completely satisfactory explanation.

The formulation of Skolem's paradox in the original paper runs as follows:

> If the axiom system of Zermelo in its precise form is consistent, then it must be possible to introduce an infinite sequence of symbols 1, 2, 3, ..., which constitute a domain B, in which all of Zermelo's axioms are valid, if these symbols are just put together in a suitable way as pairs of the form $a\epsilon b$.[4]

Skolem went on to point out that this fact only apparently led to paradoxical observations of the sort that the axioms allow us to prove the existence of large cardinals (höhere Mächtigkeiten), but that nonetheless one could (given the consistency) interpret axiomatic set theory in the natural numbers. In modern language: if set theory is consistent, then it has a denumerable model. He explained the paradox away observing that a set M in the model B which could be proved non-denumerable in set theory was also non-denumerable in the sense of B, i.e. that in B there was no bijection from M to the natural numbers in B. The enumeration of M as a set of elements of B (which followed from the denumerability of B) was not, and could not be, an object in B. Skolem's conclusion was that the notions of set theory where only *relative* to the domains (models) under consideration:

> The axiomatic founding of set theory leads to a relativity of the notions of sets, and this is inseparably connected with every systematic axiomatization.

Skolem did not view the relativity as particularly shocking or surprising, in fact he had "already communicated it to F. Bernstein in Göttingen in the winter of 1915-16".

> I believed that it was so clear that axiomatization in terms of sets was not a satisfactory ultimate foundation of mathematics that mathematicians

[4] Ist das präzisierte Zermelo'sche Axiomensystem widerspruchfrei, so muß es möglich sein, eine unendliche Reihe von Symbolen $1, 2, 3, \ldots$ so einzuführen, daß diese einen Bereich B bilden, für welchen die Zermelo'schen Axiome alle gültig sind, wenn nur diese Symbole in passender Weise zu Paaren der Form $a\epsilon b$ zusammengestellt werden.

would, for the most part, not be very much concerned with it. But in recent times I have seen to my surprise that so many mathematicians think that these axioms of set theory provide the ideal foundation for mathematics; therefore it seemed to me that the time had come for a critique. [Skolem 1923]

Skolem became a thorough going relativist who drew the ultimate (and not wholly justified) conclusion of the Skolem-Löwenheim theorem vis a vis formalized mathematics. As he put it in [Skolem 1929]:

> One recognizes here the same thing as before in the discussion of the theorem of Löwenheim, that there is no possibility to introduce something absolutely non-denumerable, but by a pure dogma.

The reception of Skolem's paradox illustrates the delay in the absorbing of new ideas in science. Fraenkel's influential "Einleitung in die Mengenlehre" [Fraenkel 1919] is a good example to trace the influence of Skolem. The second edition of 1923 mentions Skolem's paper, which had only just became available to Fraenkel, in a footnote. The paradox is referred to as "a difficulty which has so far not yet been overcome".

In the subsequent "Zehn Vorlesungen über die Grundlegung der Mengenlehre" [Fraenkel 1927], Skolem's paradox gets its own section, where it is discussed as a new, alarming attack at the axiomatic foundation of set theory. Fraenkel was not convinced of the correctness of the arguments of Skolem, he built in the *cave* "if the conclusions of Löwenheim and Skolem proceed without gaps and errors". He did not see a solution to the paradox, but was inclined to see impredicativity as a possible source of the problem. The third edition of the "Einleitung" (1928) again questions the correctness of the Skolem argument. In spite of Skolem's crystal clear exposition, Fraenkel states:

> Since neither the books have at present been closed on the antinomy, nor on the significance and possible solution so far an agreement has been reached, we will restrict ourselves to a suggestive sketch.

The remarks show that the role of logic in set theory was not quite clear to Fraenkel, no matter how much he admired Hilbert's proof theory. Apparently Skolem's arguments were beyond his expertise.

Von Neumann was quicker to grasp what was going on, his 1925 paper "An axiomatization of set theory" [Neumann 1925] shows a complete understanding of the material, not withstanding the fact that he formalized his set theory in the traditions of Hilbert's program. Von Neumann straightforwardly acknowledged the relativity phenomenon, he ended his paper with the words:

> At present we can do no more than note that we have one more reason here to entertain reservations about set theory and that for the time being no way of rehabilitating this theory is known.

Where the majority of the mathematicians followed Fraenkel's scepticism, and a few von Neumann's resignation, the first set theoretician to surpass Cantor in his own field, Ernst Zermelo, had decided that the Skolem paradox was a hoax.

2. Zermelo's Foundational Views

Between 1913 and 1929, due to "lengthy illness and isolation in a foreign country" [Zermelo 1931?], Zermelo did not publish any paper on set theory. However, the year 1929 marks the beginning of a new period of activity with invited talks on the foundations of mathematics and a series of set theoretical publications. Both his talks and the subsequent papers document an engaged fight against what he thought could become a real danger for mathematics, namely the constructive and finitistic conceptions embodied, for instance, in Brouwer's intuitionism and Skolem's first-order approach to set theory. His conviction that mathematics intrinsically is of an infinitary character had already stabilized nearly a decade earlier, as clearly stated on July 17th, 1921, in five "theses about the infinite in mathematics".[5] They contain the leading ideas of his later work. We give an English translation; the German original can be found in the Appendix.

> I) Each genuine mathematical proposition has an "infinitary" character, i.e., it refers to an *infinite* domain and has to be viewed as a combination of infinitely many "elementary" sentences.
>
> II) The infinite is neither physically nor psychologically given to us in the real world, it has to be comprehended and "posited" as an idea in the Platonic sense.
>
> III) Infinitary propositions can never be derived from finitary ones, also the "axioms" of all mathematical theories have to be infinitary, and the "consistency" of such a theory can only be proved by exhibiting a corresponding consistent system of infinitely many elementary sentences.
>
> IV) By its nature, traditional "Aristotelian" logic is finitary and, hence, not suited for a foundation of the mathematical sciences. Therefore, there is a necessity for an extended "infinitary" or "Platonic" logic which rests on some kind of infinitary "intuition" – as, e.g., with the problem of the axiom of choice –, but which, paradoxically, is refused exactly by the "intuitionists" on the ground of habit.
>
> V) Any mathematical proposition has to be conceived as a combination of (infinitely many) elementary sentences, the "ground relations", via conjunction, disjunction, and negation, and each deduction of a proposition from other propositions, in particular each "proof", is nothing but a "regrouping" of the underlying elementary sentences.

In a series of talks [Zermelo 1929A] given in Warsaw in the spring of 1929, he explicated these views more systematically. Lecture 4[6] sets the central statement: "Arithmetic – as, basically, every other mathematical discipline – essentially consists of sentences that contain infinite totalities of single propositions." The talks document that Zermelo's conception of sentences and logic are essentially of an infinitary semantical nature, mirroring a firm platonic attitude:

[5]The typewritten page has a handwritten number "42" added above the "21" in the date (a page number?), but probably not in Zermelo's handwriting. We doubt that the theses would have been formulated as late as 1942, because they clearly anticipate his Warsaw talks [Zermelo 1929A] and the Bad Elster talk [Zermelo 1932].

[6]The lectures are numbered according to their order in the manuscript.

Only if [such] a "model" exists or at least can be conceived, our system is thought of as "consistent", i.e. realizable; if it can be proved logically that the existence of such a model is impossible, the system itself is "inconsistent and contradictory".[7] [Lecture 1: *What is mathematics?*]

"Realizability" by models is exactly the basic assumption of all mathematical theories; without it also the question for the "non-contradictority" of an axiom system loses its proper meaning. For the axioms themselves do not harm each other, as long as they are not applied to one and the same (given or hypothetical) model.[8] [Lecture 2: *Disjunctive systems and the principle of the excluded middle*]

True mathematics is by its nature infinitistic and based on the assumption of infinite domains; it can even be called the "logic of the infinite".[9] [Lecture 3: *Finite and infinite domains*]

... a realization [of the assumption of infinite domains, a ground assumption of the whole of mathematics] by an explicitly given and completed model is impossible, as the infinite can nowhere tangibly be exhibited in such a way. Such an assumption can only be justified by its success, by the fact that it (alone!) has made possible the creation and development of the whole of present arithmetic which essentially is a science of the infinite.[10] (Lecture 4: *How can the assumption of the infinite be justified?*)

... the existence of the infinite as a logical postulate which ought to be the base of any "proof theory", is already made safe a priori and does not need any proof. In fact, it will not do to base the formalism again on a formalism: at one point something has really to be thought, to be posited, to be assumed. And the simplest assumption which can be made and which is sufficient for the foundation of arithmetic (just as for the whole of classical mathematics), is precisely the idea of "infinite domain"; it compulsively suggests itself to the logical-mathematical thought, and on this, in fact, our total science as it has developped historically, is built.[11]

---

[7] Nur wenn ein [solches] "Modell" existiert oder wenigstens denkbar ist, gilt unser System als "konsistent" d.h. realisierbar; ist die Existenz eines Modells als logisch unmöglich nachweisbar, so ist damit das System selber "inkonsistent" und "widerspruchsvoll".

[8] Die "Realisierbarkeit" durch Modelle ist eben die Grundvoraussetzung aller mathematischen Theorien, und ohne sie verliert auch die Frage nach der "Widerspruchsfreiheit" eines Axiomen-Systems seine eigentliche Bedeutung. Denn die Axiome selbst tun einander nichts, bevor sie nicht auf ein und dasselbe (gegebene oder hypostasierte) Modell angewendet werden.

[9] Die wahre Mathematik ist [...] ihrem Wesen nach infinitistisch und auf die Annahme unendlicher Bereiche gegründet; sie kann geradezu als die "Logik des Unendlichen" bezeichnet werden.

[10] ... unmöglich ist aber auch die Realisierung durch ein explizit gegebenes und fertig vorgelegtes Modell, weil das Unendliche eben nirgends als solches sinnfällig aufgewiesen werden kann. Rechtfertigen läßt sich eine solche Annahme lediglich durch ihren Erfolg, durch die Tatsache, daß sie (und sie allein!) die Schöpfung und Entwickelung der ganzen bisherigen Arithmetik, die eben wesentlich eine Wissenschaft des Unendlichen ist, ermöglicht hat.

[11] [...] die Existenz des Unendlichen als logisches Postulat, das jeder "Beweistheorie" zugrunde liegen müßte, [ist] bereits a priori gesichert und bedarf gar keines Beweises. Überhaupt geht es nun einmal nicht an, den Formalismus wieder auf den Formalismus zu stützen; irgend

In the time following, these views placed Zermelo outside the main foundational discussions embodied by constructivism and intuitionism on one hand and formalism on the other hand (and thus, to a certain extent, also away from Hilbert (cf. [Taylor 1993]). In [Zermelo 1931?] he describes his situation:

> These facts [the arguments of intuitionism] caused me to come back to investigations of foundational problems. I did not decide myself for a party in this proclaimed quarrel between "intuitionism" and "formalism" – actually I take this alternative to be a logically inadmissible application of the "Tertium non datur". But I thought I could help to clarify the questions under consideration, not as a "philosopher" by proclaiming "apodictic" principles which, by adding to already existing opinions, would merely augment confusion, but as a mathematician by showing objective mathematical connections – only they can provide a safe foundation for any philosophical theory.

Motivated by the discussion about his vague notion of definiteness as given in [Zermelo 1908], Zermelo published a more elaborated version in [Zermelo 1929] that sums up to second-order definability. The paper may be seen as a reaction against Fraenkel who, in the second edition of his "Einführung" (1923), had given his quasi first-order version from [Fraenkel 1922]. Besides blaming Fraenkel for the "constructive" (i.e. inductive) definition with its implicit use of the notion of finite number (which, in his opinion, should not be allowed to establish set theory but should rather be an outcome of it), Zermelo states that any attempt to make the notion precise by means of logic would be as unsuccessful as in 1908:

> A generally accepted "mathematical logic" to which I could have appealed, did not exist at that time, as it doesn't today where each foundational researcher has his own system of logic (seine eigene Logistik).

It is not quite clear whether Zermelo really knew Skolem's precise treatment in [Skolem 1923] or whether he was really aware of its scope when he gave his second-order definition, because he does not mention Skolem, whereas later, Skolem is *the* crystallization point of his criticism. Also Skolem, in his sagacious reply [Skolem 1930] hints at this question ("... he seems not to know my Helsingfors talk."). In any case, through Skolem's reply Zermelo was fully informed and clearly confronted with an opinion sharply contradicting his views as laid down in the documents cited above. Of course, Zermelo did not object to axiomatizations in general – the point of his critique was just the underlying

einmal muß doch wirklich gedacht, muß etwas gesetzt, etwas angenommen werden. Und die einfachste Annahme, die gemacht werden kann und die zur Begründung der Arithmetik (wie auch der gesamten klassischen Mathematik) ausreicht, ist eben jene Idee der "unendlichen Bereiche", die sich dem logisch-mathematischen Denken geradezu zwangsmäßig aufdrängt, und auf die auch tatsächlich unsere gesamte Wissenschaft, so wie sie sich historisch entwickelt hat, aufgebaut ist.

language, where, in particular, any first-order approach was doomed to fail because the richness of mathematics could not be caught in a finitary way.

His answer to the finitary challenge was twofold: In [Zermelo 1930] he drops any restriction in the axiom of separation, and in [Zermelo 1932] he develops a system of infinitary logic.[12]

In the first paper he states that the separating property "may be totally *arbitrary* ... and all conclusions that have been drawn by limiting oneself to a special class [of properties] fall away for the point of view taken here" and moreover, "to each part of a set there corresponds a set which contains all elements of this part".[13] In his 1908 paper Zermelo had introduced definiteness in order to be on the safe side with respect to the paradoxes of Richard and Russell. The cancellation of any restriction is possible now as there is no longer any danger of getting involved in the paradoxes: Each "Unmenge" of a model of set theory becomes a set on the next level in the cumulative hierarchy of models:

> The "ultrafinite antinomies of set theory" ... are based only on a confusion of *set theory itself* as resting on non-categorical axioms with the individual *models* that represent it: an ultrafinite non- or superset of one model becomes a valid "set" with cardinal number and ordinal type in the next higher model and thus forms the foundation stone for the construction of the new domain.[14]

It is not clear to us, whether Zermelo has really given up the limitation of separation by definiteness. In the paper the connection of unlimited separation with the notion of definiteness and with the criticism in [Skolem 1930] is expressis verbis transferred to a later discussion. However, there seems to be evidence that the concept of the cumulative hierarchy with its free formation of subsets looked quite convincing to him:

> The diametrically opposed tendencies of the thinking spirit, the idea of creative *progress* and the idea of summarizing *conclusion* find their symbolic representation and their symbolic reconciliation in the transfinite series of numbers that, founded on the notion of well-ordering, by its unlimited progression does not have a true end, but has relative stops, namely those "limit numbers" that separate the models of higher type from the ones of lower type. [15]

[12]Of course, it would be inadequate to consider these papers mainly in the light of the present discussion. They both mark essential developments in logic: [Zermelo 1932] for the first time considers members of the cumulative hierarchy as models of set theory (cf. [Kanamori 1994] 19ff.); [Zermelo 1932] contains the first definition of infinitary languages; they were to be developed again only in the late fifties in the group around Tarski.

[13]For a discussion of other ways of reading these remarks, e.g. as making precise definiteness by second-order definability, see [Taylor 1993].

[14]Die "ultrafiniten Antinomien der Mengenlehre" [...] beruhen lediglich auf einer Verwechselung der durch ihre Axiome nicht-kategorisch bestimmen *Mengenlehre selbst* mit den einzelnen sie darstellenden *Modellen*: was in einem Modelle als "ultrafinite Un- oder Übermenge" erscheint, ist im nächst höheren bereits eine vollgültige "Menge" mit Kardinalzahl und Ordnungstypus und bildet selbst den Grundstein zum Aufbau des neuen Bereiches.

[15]Die beiden polar entgegengesetzten Tendenzen des denkenden Geistes, die Idee des

Even more, he used the loss of uniqueness that went together with the cumulative series of models, as an argument against Skolem, for in [Zermelo 1931?] he comments on his results:

> I thus ended up with some kind of "set-theoretical relativism" that is, however, basically different from Skolem's "relativism"; the latter one even relativizes the notions of "subset" and "cardinality": Skolem wants to restrict the formation of subsets to special classes of defining functions, whereas I, according to the true spirit of set theory, admit *free* separation and postulate the existence of subsets in whatever a way they are formed. According to Skolem the *whole* of set theory should be representable in a *countable* model, and for him already, say, the problem of the power of the continuum loses its genuine meaning.

In an appendix of his report he once more lists the main points, in the forth point expressively stating that he "avoid[s] Skolem's "relativism" by admitting *free* formations of subsets *without* limitations of definiteness".

In his second answer to the finitary challenge, [Zermelo 1932], with a continuation in [Zermelo 1935], he works out the idea of infinitistic formulae and infinitistic proofs as already present in the theses of 1921 and permeating the whole of his Warsaw lectures from 1929. The introduction contains his infinitary credo:

> When starting from the assumption that all mathematical notions and theorems should be representable by a *fixed finite system of signs*, already the arithmetical continuum inevitably leads to the well-known "Paradox of Richard"; seemingly settled and buried, this paradox now has found a happy resurrection as "Skolemism", the doctrine according to which every mathematical theory, including set theory, can be realized by a *countable* model. It is well-known that inconsistent premises can prove anything one wants: however, even he strangest consequences that Skolem and others have drawn from their basic assumption, for instance the relativity of the notion of subset or equicardinality, still seem to be unsufficient to raise doubts about a doctrine that, for various people, already has won the power of a *dogma* that is beyond all criticism. However, a sound "metamathematics", a true "logic of the infinite", will only be possible by thoroughly *turning away* from the assumption that I have described above and which I would name the "finitary prejudice". Generally speaking, it is not the "formations of signs" that – according to the opinion of various people – form the true subject of mathematics, but the conceptual and ideal relations between the elements of infinite varieties that are set in a conceptual way, and our notational systems are only *defective* means of our finite intelligence to at least gradually approach and dominate the infinite that we cannot directly and "intuitively" survey or perceive.

schöpferischen *Fortschrittes* und die des zusammenfassenden *Abschlusses* [...] finden ihre symbolische Darstellung und ihre symbolische Versöhnung in der auf den Begriff der Wohlordnung gegründeten transfiniten Zahlenreihe, die in ihrem schrankenlosen Fortschreiten keinen wahren Abschluß wohl aber relative Haltepunkte besitzt, eben jene "Grenzzahlen", welche die höheren von den niederen Modelltypen scheiden.

And he continues that he will develop a kind of mathematical logic that, *free* of the finitistic prejudice, should allow one to build up the whole of mathematics without *arbitrary prohibitions* and *limitations*.

Besides the manuscript of the published version, The Nachlass contains a manuscript with the title "Über mathematische Systeme und die Logik des Unendlichen. Vortrag gehalten auf der Mathematiker-Tagung zu Bad Elster September 1931", which has a similar contents, but the introduction of the published version that we have just quoted, is missing. On the meeting also Gödel had given a talk, presenting his incompleteness results. Apparently it were these results and, in particular, the underlying way of seeing mathematics as a finitary formal system that caused Zermelo to add the general remarks. His first letter to Gödel, less than a week after the meeting, says that, when writing the final version of his talk, he *had* to refer to Gödel's talk. Another reason for writing the letter may be seen in the fact that Zermelo believed to have detected a gap in Gödel's argument[16] which he had located as to rest on "the (erroneous) assumption that each mathematically definable notion could be expressed by a "finite combination of signs" (according to a *fixed* system)". With the possibility of winning Gödel as a comrade, he continued that "in particular your proof, if interpreted in the right way [that is, as detecting insufficiencies of any finitary approach to mathematics], could ... render an essential service for truth". The second letter to Gödel from October 29, 1931 (cf. [Grattan-Guinness 1979]) ends with a statement about proofs and provability that once more very distinctly sets off the formal approach against his own infinitary platonic conception of mathematics.

> What does one understand by a proof? In general, a proof is understood as a system of sentences that, when accepting the premises, yields the validity of the assertion as being *reasonable*. And there remains only the question of what may be "reasonable". In any case – as you are showing yourself – *not only* the propositions of some finitary scheme that, also in your case, may always be *extended*. So with this respect we are of the same opinion; however, I a priori accept a *more general* scheme that *need* not be extended. And in *this* system, really *all* propositions are decidable.

3. Zermelo's attempted Refutation of the Skolem Paradox

In [Zermelo 1931?] Zermelo lists five questions he planned to work upon. Items 3 and 4 say: "On the set theoretical relativism with Skolem and me and its significance for the continuum problem" and "On mathematics as a "logic of the infinite" and the impossibility of a "finitistic mathematics"". What could be more convincing with respect to this program than to really refute Skolem's paradox? The idea was not new: In a letter of May 25, 1930, addressed to

[16]cf. [Dawson 1997] or [Grattan-Guinness 1979].

"a colleague",[17] Zermelo offers to give a talk in Hamburg on his results in [Zermelo 1930]; he writes:

> ...by "relativizing" the notion of set in this way, I believe to be able to refute Skolem's "relativism" that would like to represent the whole of set theory in a *countable* model. It simply is *impossible* to give *all* sets in a constructive way ...and any theory, founded on this assumption, would by no means be a theory of sets."[18]

On October 4, 1937, under the title "Relativism in set theory and the so-called Theorem of Skolem", he wrote down what he thought would be such a refutation:

> In set theory one usually defines the "continuum" as the set P of all subsets M_1 of a "countable" set M. However, is this definition really *unambiguous*, is the notion of "all" subsets not too uncertain? Couldn't there be *different* degrees of this totality and, hence, different "models" of the continuum; even more, could this continuum, regardless of its formal properties (according to "Skolem's Theorem") be representable by a *countable* model? In this case it should be possible to imbed such a "meager" continuum K' in a "fat" continuum K in such a way that all formal properties carry over to K.
>
> We consider a *countable* set M and a subtotality K' of the totality K of *all* subsets N (in the sense of the "fatter" continuum) with the following property:
>
> 1. If R is a subset of K' then both the union set S_R belonging to R and its intersection D_R are elements of K! In other words: The sum and the intersection of arbitrarily many (even ∞ many) sets from K' are elements of K' again, the elements of K' form a "ring of sets".
> 2. Each element m of M is an element of at least one of the sets in K'.
>
> Now, let N_m be the *intersection* of all sets N from K' which contain m. Hence, N_m is an element of K', and two such intersections N_a and N_b are either identical or disjoint. Thus the total set M is partitioned into a sum of mutually disjoint parts N_m, each of which has to be finite or denumerable. Moreover, also *each* set N from K' is partitioned into such parts N_m and, by 1), each union, each sum of such N_m has to be an element of K' again.
>
> Hence, if we denote the set of all (mutually distinct) N_a by T, then each element of K' uniquely corresponds to a subset of T, and K' is equivalent[19] to the set $\mathbf{U}T$ of all subsets of T. Now the set of parts into which a *denumerable* set M can be partitioned is either *finite or denumerable* itself, that is, equivalent to M, and hence, the pseudo-continuum K' is either finite itself or equivalent to $K' \sim \mathbf{U}T \sim \mathbf{U}M = K$, that is, equivalent

[17] We are grateful to V. Remmert for the suggestion that the recipient was probably Emil Artin.

[18] Indem ich so den Mengenbegriff " relativiere", glaube ich andererseits den Skolemschen "Relativismus", der die ganze Mengenlehre in einem *abzählbaren* Modell darstellen möchte, widerlegen zu können. Es können eben nicht alle Mengen konstruktiv gegeben sein [...] und eine auf diese Annahme gegründete Theorie wäre überhaupt keine Mengenlehre mehr.

[19] 'equivalent' means 'gleichmächtig', 'equipollent'

> to the original (uncountable) continuum K. Therefore it is impossible to represent the continuum in a *denumerable* model, it then would have to be *finite*. "Skolem's Theorem" thus leads to the interesting consequence that *infinite* sets can be realised in finite models – a consequence that would not be more paradox than many other consequences already obtained from this nice theorem. In this way also the ideal aim of intuitionism, the abolition of the infinite from mathematics, would have been brought nearer to realization together with the aim of "formalism" that, as we know, aims at the proof of consistency. For, as we know, from absurd premises one can prove *anything*. Hence also the consistency of an arbitrary system of sentences.

The handwritten note never reached its final stage. In the present form there are some minor inaccuracies that Zermelo doubtlessly would have spotted and repaired. The argument is clever, and it would probably have confused most readers, it certainly confused Zermelo himself. For us it easy to see where the argument goes wrong, but in the thirties there was little or no experience with models of set theory. It is likely that Zermelo was not satisfied with the refutation of the Skolem phenomenon, otherwise he would probably have circulated (if not published) the note.

The modern reader, when considering the correctness of the argument would think of a countable elementary submodel V' of a standard model V of ZF, he would take M to be ω, and pick for K and K' the respective continuums in V and V'. Now Zermelo's proof makes use of arbitrary intersections and unions of subsets of K', whereas for the argument 'arbitrary' has to mean 'in the sense of V'. This closure property, however, is not available. So here Zermelo's proof is in error. Should we allow Zermelo's assumption, then we know that elementary equivalence cannot apply (what Zermelo calls "all formal properties carry over"), and all that remains is a proof that all powersets of denumerable sets are equivalent.

How could this happen? Surely, Zermelo would have been able to mathematically understand, say, Skolem's proof of the Löwenheim-Skolem Theorem. However, he seems to have been blocked to view this theorem as a purely mathematical statement and, instead, was caught in the special case of a first-order axiomatization of set theory, a system that – as we have seen – totally contradicted his understanding of set theory and strongly evoked his epistemological resistance. So closure of K' under *all* unions was self-evident for him, and considering unions in the sense of some inadequate model of set theory meant a mode of thinking he refused to perform. In his own terms from [Zermelo 1930] : Faced with "restriction and mutilation", he firmly had decided himself for "unfolding and enrichment".

Apparently Zermelo has not discussed his argument with others. In a concrete sense the manuscript was written down at a time of growing isolation and retreat: Two and a half years earlier, having been reported to the university by a colleague in the mathematics department, he had been given only the choice of either losing his honorary professorship or opening his lectures with the Hitler salute, and he had chosen the first possibility.

So there are two ways of reading the note on "Relativism". The first reading would be the modern first-order set theoretical one: Zermelo simply committed an error; the second, more charitable, one would consider the note as a demonstration that in Zermelo's version of set theory no Skolem phenomenon could occur.

Note that this is similar to what had happened before in the case of Gödel's incompleteness results. As indicated at the end of the second section[20] also here Zermelo's reaction is characterized by a certain technical weakness coming along with a firm epistemological conviction. And again, we are offered an alternative way of judging this. As we believe, in both cases, an "either–or" would be inadequate. Rather, the choice should be an "as well–as", where, in addition, we might add that there seems to be a mutual interchange: The epistemological preoccupancy may have prevented his technical attention, and the missing technical penetration may have resulted in an even stronger perseverence on his own views.

One remark may be added before we end our paper. This is not the first time that a too inflexible philosophical position is a danger for a balanced view of foundational matters. By a curious coincidence, Skolem himself was a victim of the same phenomenon: In 1922 he basically proved the completeness theorem for predicate logic (cf. Gödel's remark in [van Heijenoort 1967], p. 510). It seems that his finitistic conviction withheld him from recognizing the full set theoretic–semantic significance of his results.

Appendix

ZERMELO 17. Juli 1921

Thesen über das Unendliche in der Mathematik

I) Jeder echte mathematische Satz hat "infinitären" Charakter, d.h. er bezieht sich auf einen *unendlichen* Bereich und ist als eine Zusammenfassung von unendlich vielen "Elementarsätzen" aufzufassen.

II) Das Unendliche ist uns in der Wirklichkeit weder physisch noch psychisch gegeben, es muß als "Idee" im Platonischen Sinne erfaßt und "gesetzt" werden.

III) Da aus finitären Sätzen niemals infinitäre abgeleitet werden können, so müssen auch die "Axiome" jeder mathematischen Theorie infinitär sein, und die "Widerspruchslosigkeit" einer solchen Theorie kann nicht anders "bewiesen" werden als durch

[20] See, moreover, the literature cited there.

Aufweisung eines entsprechenden widerspruchsfreien Systems von unendlich vielen Elementarsätzen.

IV) Die herkömmliche "Aristotelische" Logik ist ihrer Natur nach finitär und daher ungeeignet zur Begründung der mathematischen Wissenschaft. Es ergibt sich daraus die Notwendigkeit einer erweiterten "infinitären" oder "Platonischen" Logik, die auf einer Art infinitärer "Anschauung" beruht - wie z.B. in der Frage des "Auswahlaxioms" –, aber paradoxerweise gerade von den "Intuitionisten" aus Gewohnheitsgründen abgelehnt wird.

V) Jeder mathematische Satz ist aufzufassen als eine Zusammenfassung von (unendlich vielen) Elementarsätzen, den "Grundrelationen", durch Konjunktion, Disjunktion und Negation, und jede Ableitung eines Satzes aus anderen Sätzen, insbesondere jeder "Beweis" ist nichts anderes als eine "Umgruppierung" der zugrunde liegenden Elementarsätze.

Der Relativismus in der Mengenlehre und der sogenannte Skolem'sche Satz

Das "Kontinuum" wird in der Mengenlehre gewöhnlich definiert durch die Menge P *aller* Untermengen M_1 einer "abzählbaren" Menge M. Aber ist diese Definition auch *eindeutig*, ist der Begriff "aller" Untermengen nicht zu unbestimmt? Könnte es nicht *verschiedene* Grade dieser Allheit und damit verschiedene "Modelle" des Kontinuums geben, könnte nicht vielleicht dieses Kontinuum unbeschadet seiner formalen Eigenschaften sogar (entsprechend dem "Skolemschen Satze") durch ein *abzählbares* Modell dargestellt werden? Dann müßte es möglich sein, ein solches "mageres" Kontinuum K' in ein "fetteres" K so einzubauen, daß sich alle formalen Eigenschaften auch auf dieses übertragen lassen.

Wir betrachten eine *abzählbare* Menge M und aus der Gesamtheit K *aller* ihrer Untermengen N (im Sinne des "fetteren" Kontinuums) eine Teilgesamtheit K' von folgender Eigenschaft:

1. Ist R eine Untermenge von K', so sind sowohl die zu R gehörige Vereinigungsmenge S_R wie ihr Durchschnitt D_R wieder Elemente von K'. M.a.W. Summe und Durchschnitt beliebig vieler (auch ∞ vieler) Mengen aus K' gehören wieder zu K', die Elemente von K' bilden einen "Mengenring".
2. Jedes Element m von M ist in mindestens einer Menge aus K' als Element enthalten.

Dann sei N_m der *Durchschnitt* aller Mengen N aus K', welche m enthalten, also selbst ein Element von K', und zwei solche Durchschnitte N_0 und N_1 sind entweder identisch oder elementenfremd. Dadurch wird die Gesamtmenge M zerlegt in eine Summe von elementenfremden Bestandteilen N_m, welche einzeln entweder endlich oder abzählbar sein müssen. Ferner wird *jede* Menge N aus K' gleichfalls in solche Teile N_m zerspalten und jede Vereinigung, jede Summe solcher N_m muss nach 1) wieder ein Element von K' sein.

Bezeichnet man also mit T die Menge aller (von einander verschiedenen) M_a[21], so entspricht jedes Element von K' ein-eindeutig einer Untermenge von T und K' ist äquivalent der Menge $\mathbf{U}T$ aller Untermengen von T. Nun ist die Menge der Teile, in die eine *abzählbare* Menge M zerspalten werden kann, entweder *endlich oder* selbst *abzählbar*, d.h. der Menge M äquivalent, und das Pseudo-Kontinuum K' daher entweder selbst endlich oder äquivalent $K' \sim \mathbf{U}T \sim \mathbf{U}M = K$, d.h. äquivalent dem ursprünglichen (nicht abzählbaren) Kontinuum K. Es ist daher unmöglich, das Kontinuum in einem *abzählbaren* Modell darzustellen, es müsste dann *endlich* sein. Der "Skolem'sche Satz" führt also zu der interessanten Folgerung, daß sich *unendliche* Mengen in endlichen Modellen realisieren lassen – eine Folgerung, die nicht paradoxer wäre als manche andere aus diesem schönen Satze bereits gezogene Konsequenzen. Hiermit wäre denn auch das Ideal des "Intuitionismus", die Abschaffung des Unendlichen aus der Mathematik, der Verwirklichung nahe gebracht zugleich mit dem des "Formalismus", der bekanntlich den Beweis der Widerspruchsfreiheit erstrebt. Denn aus absurden Prämissen kann man bekanntlich *alles* beweisen. Also auch die Widerspruchsfreiheit eines beliebigen Satzsystems.

References

[Dawson 1997] J.W. Dawson. *Logical Dilemmas. The Life and Work of Kurt Gödel.* A.K. Peters, Wellesley, Mass.

[Dawson 1985] J.W.D. Dawson. Completing the Gödel-Zermelo correspondence. *Historia Mathematica*, 12:66–70.

[Fraenkel 1919] A. Fraenkel. *Einleitung in die Mengenlehre.* Springer, Berlin.

[Fraenkel 1922] A. Fraenkel. Über den Begriff "definit" und die Unabhängigkeit des Auswahlaxioms . *Die Preussische Akademie der Wissenschaften. Sitzungsberichte. Physikalische-Mathematische Klasse.*:253–257.

[Fraenkel 1927] A. Fraenkel. *Zehn Vorlesungen über die Grundlegung der Mengenlehre.* Teubner, Leipzig.

[Grattan-Guinness 1979] I. Grattan-Guinness. In memoriam Kurt Gödel: His 1931 Correspondence with Zermelo on his incompletability theorem. *Historia Mathematica*, 6:294–304.

[van Heijenoort 1967] J. van Heijenoort. *From Frege to Gödel* Harvard University Press, Cambridge, Massachusetts.

[Kanamori 1994] A. Kanamori. *The Higher Infinite. Large Cardinals in Set Theory from Their Beginnings.* Springer, Berlin.

[Moore 1980] G.H. Moore. Beyond first-order logic: The historical interplay between mathematical logic and axiomatic set theory. *Hist. and Phil. of Logic*, 1:95–137.

[Neumann 1925] J. von Neumann. Eine Axiomatisierung der Mengenlehre. *J Reine Angew Math*, 154:219–240.

[21] M_a should be replaced by N_a.

[Skolem 1923] Th. Skolem. Einige Bemerkungen zur axiomatischen Begründung der Mengenlehre. In *Proc. 5th Scand. Math. Congress, Helsinki 1922*, pages 217–232,. Akademiska Bokhandeln, Helsinki.

[Skolem 1929] Th. Skolem. Über einige Grundlagenfragen der Mathematik. *Norsk Vid.-Akad. Oslo. Mat.-nat. Klasse*, 4:1–29.

[Skolem 1930] Th. Skolem. Über die Grundlagendiskussionen in der Mathematik. In *Proc. 7th Scand. Math. Congress, Oslo 1929*, pages 3–21, Broggers. Oslo.

[Taylor 1993] R.G. Taylor. Zermelo, reductionism and the philosophy of mathematics. *Notre Dame Journal of Formal Logic*, 34:539–563.

[Zermelo 1908] E. Zermelo. Neuer Beweis für die Möglichkeit einer Wohlordnung. *Math Ann*, 65:107–128.

[Zermelo 1929] E. Zermelo. Über den Begriff der Definitheit in der Axiomatik. *Fund. Math.*, 14:339–344, 1929.

[Zermelo 1929A] E. Zermelo. Vortrags-Themata für Warschau 1929. Manuscript in the Nachlass. Manuscript in the Nachlass. Partly reproduced in [Moore 1980]. .

[Zermelo 1930] E. Zermelo. Über Grenzzahlen und Mengenbereiche. Neue Untersuchungen über die Grundlagen der Mengenlehre. *Fund. Math.*, 16:29–47, 1930.

[Zermelo 1931?] E. Zermelo. Bericht an die Notgemeinschaft der Deutschen Wissenschaft über meine Forschungen betreffend die Grundlagen der Mathematik. Manuscript in the Nachlass. Reproduced in [Moore 1980].

[Zermelo 1932] E. Zermelo. Über Stufen der Quantifikation und die Logik des Unendlichen. *Jahresber. Dtsch. Math. Ver.*, 41:85–92.

[Zermelo 1935] E. Zermelo. Grundlagen einer allgemeinen Theorie der mathematischen Satzsysteme (Erste Mitteilung). *Fund. Math.*, 25:136–146.