Notes on Zermelo's axioms for set theory

Randall Holmes

September 24, 2020

I'm going to organize these point by point using the section numbers in Zermelo's paper, with some intermissions.

1. He introduces a domain **B** of all objects. He says that some of these objects are *sets* (primitive notion). He introduces equality as a primitive notion.

2. In section 2, he introduces membership $a \in b$ as a primitive relation. He tells us that if $a \in b$, b must be a set (and telegraphs that there can be a set with no elements, but there can be only one: this follows from a later axiom). In any event what he says here is enough to see that if there is any object which is not a set, it has no elements.

   It is a feature of most modern treatments of foundations of mathematics based on set theory that all objects are sets. This isn't a natural assumption philosophically, though philosophers have gamely played along with it, and it is nice to see that in this source text this assumption is not made.

3. In section 3, he defines the subset relation: $M \subseteq N$ iff M and N are sets (important) and for all x, if x is an element of M then x is an element of N. Notice that for any x which is not a set and any y, it is true that any z which belongs to x (there aren't any) belongs to y, but we do not say $x \subseteq y$. He also defines disjointness of sets.

4. In section 4, he introduces the idea of an assertion ϕ being "definite", meaning that it can be decided whether it is true or false. This is important to him, and more in a philosophical than a mathematical way. A propositional function $P(x)$ of an object x ranging over a

class K is definite if $P(x)$ is definite for each element of the class K. Question: what does he mean by "class"? Is $\mathbf{B}$ a class in this sense? He does know that $\mathbf{B}$ is not a set (see section 10). He tells us that $a \in b$ is always definite, and that $M \subseteq N$ is always definite (because the propositional function $x \in N$ of x is definite for each $x \in M$, I interpret him as saying).

He now introduces some axioms:

Axiom I: If $M \subseteq N$ and $N \subseteq M$ then $M = N$. Sets with the same elements are the same. This is called the axiom of extensionality.

informal definition: He introduces the notation $\{a, b, c, \ldots, r\}$ for a finite set whose elements are exactly $a, b, c, \ldots, r$. A motto of mine: whenever a mathematician introduces those rows of dots, he is cheating.

Axiom II: There is a set 0 with no elements (why does he feel constrained to call it "fictitious"?); for any object a the set $\{a\}$ which has a as an element and no other element exists; for any two objects a, b, the set $\{a, b\}$ which has a as an element, has b as an element and has no other elements exists. This is called the axiom of elementary sets.

5. In section 5 he observes that the sets 0, $\{a\}$ and $\{a, b\}$ are uniquely determined by axiom I. He also notes that $a = b$ is always "definite" because it is equivalent to $a \in \{b\}$. I was going to say in protest that it is definite because it is equivalent to $a \subseteq b \wedge b \subseteq a$, but that is actually inadequate: that only shows that equality of sets is definite. The question of why Zermelo wants to talk about definiteness of assertions is really interesting. Does he ever say anything that isn't definite?

   The notation $\emptyset$ for the empty set is more usual now, but we will stick with 0.

6. In section 6 he raises a really interesting issue in philosophy of set theory as it were in passing. He notes that $0 \subseteq M$ and $M \subseteq M$ are always true: he defines a "part" of a set M as a subset of M other than 0 or M. We might think of an object as a part of itself, and prefer to say "proper part" here. We can feel with Zermelo discomfort with the idea that all sets have a common part 0. But in any event, the

philosophical point to make is that the elements of a set are not parts of the set (whatever a part of a set may be taken to be). The relation of part to whole, however it is defined, must be transitive, and $a \in \{a, b\}$, $\{a, b\} \in \{\{a, b\}\}$ but $a \notin \{\{a, b\}\}$. A set does not have its elements as its parts but its subsets; it does have (or can be understood as having) its singleton subsets as atomic parts correlated with its elements. David Lewis wrote a whole book about this, *Parts of Classes*, which I quite recommend as reading.

Now (still in section 6) we get

Axiom III: Whenever the propositional function $P(x)$ is definite for each element of a set M, there is M_P such that $M_P \subseteq M$ and for every x, $x \in M_P$ if and only if $x \in M$ and $P(x)$. This is called the Axiom of Separation.

We introduce the notation $\{x \in M : P(x)\}$ for the set M_P.

This axiom gives the kind of ability to define objects correlated with properties which Frege wanted in his Axiom V governing "courses of values". But the restriction of this formation of objects from properties to properties of elements of a previously given set seems to make this workable without contradiction. Moreover, nothing is being given up in terms of actual mathematical practice: we do not construct sets of mathematical interest by considering properties of all objects taken indiscriminately, but by considering properties of objects of a particular sort.

Zermelo talks about the importance of ensuring that the property $P(x)$ is "definite". I am quite interested in what he thinks this notion is doing for him.

7. In section 7, Zermelo notes that for $M_1 \subseteq M$ (and in fact for any M_1) we can define $M \setminus M_1$ as $\{x \in M : x \notin M_1\}$, Zermelo calls this the complement of M_1; it would more usually be called the complement of M_1 relative to M. This is provided by Axiom III.

8. In section 8, Zermelo introduces the intersection $[M, N]$ of sets M and N, provided by axiom III as $\{x \in M : x \in N\}$. This is more usually written $M \cap N$ now. The ability to write the assertion that M and N are disjoint as $[M, N] = 0$ is worth noting.

9. In section 9 he extends intersections first to intersections $[M, N, R, \ldots]$ of a list of given sets and then to the notion $\bigcap T$ of the intersection of all elements of a nonempty set T. His analysis of this is interesting. He says that by Axiom III for a given set T and each a we can define a subset $T_a = \{t \in T : a \in T\}$ of all elements of T which contain a. It is then a definite question for each a whether $T_a = T$ (that is, whether every element of T contains a) and so by Axiom III there is a set $\bigcap T = \{a \in t : T_a = T\}$, where t is any element of T (it does not matter which one you use). A more usual definition would be $\{a \in t : (\forall u \in T : a \in u)\}$: but Zermelo in his formulation avoids using a quantifier.

10. In section 10 we see the form which the Russell paradox argument takes in this theory. It does not lead to paradox.

    Theorem: For each set M, there is a set M_0 such that $M_0 \subseteq M$ and $M_0 \notin M$.

    Proof: It is definite for each $x \in M$ whether $x \in x$ (and Zermelo takes pains to note that nothing in his system prevents $x \in x$ from being true for some x).

    Let $M_0 = \{x \in M : x \notin x\}$. Clearly $M_0 \subseteq M$.

    Now either $M_0 \in M_0$ or not. If $M_0 \in M_0$ then M_0 contains an x such that $x \in x$, which is incompatible with its definition. So $M_0 \notin M_0$. But then $M_0 \in M$ is impossible, as if we had $M_0 \in M$ it would meet the conditions to belong to M_0, and we have already shown that it cannot meet these condiitions. So $M_0 \notin M$ as desired.

    Zermelo then observes that it follows from this that we cannot have a set which contains every object, and so (interestingly) he observes that the domain **B** cannot be a set. It is philosophically very interesting that he makes this observation: one could further ask what sort of thing **B** is... and this is a natural opening to the later idea of proper classes.

    Still in section 10, Zermelo introduces two more axioms.

    Axiom IV: To every set A there corresponds a set $\mathcal{P}(A)$, the power set of A, whose elements are exactly the subsets of A. (Axiom of power set)

Axiom V: To every set A, there corresponds a set $\bigcup A$, the union of A, whose elements are exactly the elements of the elements of A. (Axiom of union)

side comments of ours about these axioms: It is very interesting that the union $\bigcup A$ has to be provided by axiom where the related set $\bigcap A$ is provided already by Axiom III.

An interesting point about the power set, which is related to the intellectual origins of Russell's paradox and the theorem of section 10, is the theorem of Cantor that the power set of A, even for A infinite, is larger in size than the set A.

We give an informal argument for this. We say that two sets A and B are the same size if there is a bijection from A to B, that is, a function F from A to B which is one to one and onto B. The set $\mathcal{P}(A)$ is at least as large as A (each element a of A corresponds to $\{a\} \in \mathcal{P}(A)$). Suppose that A is at least as large as $\mathcal{P}(A)$, that is, there is a one-to-one map f from $\mathcal{P}(A)$ to a subset of A. This map has an inverse f^{-1} defined for some but not all elements of A. For the sake of argument we extend $f^{-1}(a)$ to be 0 for a in A which is not in the range of f. Now define R as $\{a \in A : a \notin f^{-1}(a)\}$. We ask whether $f(R)$, a well-defined element of A by hypothesis, is an element of R: this will be true if and only if $f(R) \in A$ and $f(R) \notin f^{-1}(f(R))$. Now $f(R) \in A$ is true and $f^{-1}(f(R)) = R$, so we have $f(R)$ an element of R if and only if $f(R)$ is not an element of R, which is a contradiction: there can be no such f.

This is not a proof in Zermelo's system, because we do not yet know how to talk about functions like f in Zermelo's system, but shortly we will.

11. In section 11 we define binary unions of sets. We define $M + N$ (usually now written $M \cup N$) as $\bigcup\{M, N\}$. More generally we define $M + N + R + \ldots$ as $\bigcup\{M, N, R, \ldots\}$. Some algebraic laws such as $M + 0 = M + M = M$ are noted.

12. In section 12 it is noted that the commutative and associative laws $M + N = M + N$ and $(M + N) + R = M + (N + R)$ hold for union of sets. Further, distributive laws $[M + N, R] = [M, R] + [N, R]$ and $[M, N] + R = [M + R, N + R]$ hold. Zermelo notes that these theorems are proved by Axiom I and logic.

13. If M is a set different from 0 and $a \in M$, it is definite whether $M = \{a\}$. It is definite whether a set has one element or not. This is Zermelo's language: here is my own argument: for any set M. it is definite whether there is an element a of M such that every element x of M is equal to a. If there is such an element a of M, then $M = \{a\}$, and if there is not, M is not a singleton.

We say that a set T is a pairwise disjoint collection of sets or a partition if for each pair A, B of distinct elements of T we have that A and B are disjoint.

Let T be a pairwise disjoint collection of sets. We define $\prod T$ as the collection of all subsets X of $\bigcup T$ such that for each $A \in T$, $A \cap X$ has exactly one element. A product $\prod\{M, N\}$ is written MN and a product $\prod\{M, N, R, \ldots\}$ is written $MNR \ldots$.

Before stating the next axiom, it is worth noting the relationship between the addition and multiplication operations now defined on sets and addition and multiplication of numbers. If M has m elements and N has n elements and M, N are disjoint, then $M + N = M \cup N$ has $m+n$ elements and MN has mn elements. Similar statements are true for sums and products of finite collections of disjoint sets.

We would like it to be the case that a product of infinitely many nonempty sets has to be nonempty, and that is what the next axiom says. The final axiom provides us with something no previous axiom has done, an example of an infinite set.

Axiom VI: If T is a pairwise disjoint collection of sets and $0 \notin T$, then $\prod T$ is nonempty. In other words, for any pairwise disjoint collection P of nonempty sets, we can find a set which contains exactly one element of P: we can choose one element from each set and collect them. This is called the Axiom of Choice.

A Holmes side remark: the Axiom of Choice allows us to choose an element from each element of a possible infinite pairwise disjoint set T. An evil fact is that Zermelo's theory allows us to choose for each element t of a set T a set *not* in T without using Choice and collect the resulting objects into a set. The collection of all sets $t_0 = \{x \in t : x \notin x\}$ does not belong to t and does belong to the power set of t and so to the power set of $\bigcup T$. So we can define the "negative choice set" N_T as $\{u \in \mathcal{P}(\bigcup T) : (\exists t \in T : u = \{x \in t : t \not\in t\})\}$: we can choose an element of the power set of the union of T which does not belong to t for each $t \in T$ and collect these into a set, without using Axiom VI. If T is a partition (a pairwise disjoint collection of nonempty sets) and contains no t which is its own singleton set, then a distinct non-element will be chosen for each element of T. Of course, there is less to this than it seems: if we have $(\forall x : x \not\in x)$, then $t_0 = t$ and $N_T = T$. But it is still fun.

Axiom VII: There is a set Z such that $0 \in Z$ and for each element a of Z, $\{a\}$ is also an element of Z. (axiom of infinity).

Observe that Axiom II already gives us infinitely many distinct objects, $0, \{0\}, \{\{0\}\}, \ldots$, but nothing before axiom VII allows us to construct an infinite set.

what if everything is a finite set? Nothing under this heading is in Zermelo: these are entirely my modern comments.

I actually pause to investigate this. I give an alternative collection of axioms:

A: same as Axiom I: sets with the same elements are the same.

B: The empty set 0 exists. For each object a, $\{a\}$ exists.

C: for each nonempty set A and object b, $A + \{b\} = A \cup \{b\}$ exists.
We note that it would work simply to assert that the empty set 0 exists and for any set x and object y, $x \cup \{y\}$ exists.

D: for each propositional function $P(x)$ such that $P(0)$, for each object a, $P(a) \to P(\{a\})$, and for each set A and object b, $P(A) \wedge P(b) \to P(A \cup \{b\})$ is true, we have that $P(x)$ is true for all x. What Axiom D is saying is that we allow only those objects which are either 0, or the singleton of a previously constructed object, or of the form $A \cup \{b\}$ where A and b are already constructed: we are only allowing sets built by finite listing from previously given sets.
In line with the suggestion under Axiom C, this could be simplified to say, for any propositional function $P(x)$, that if (1) $P(0)$ is true and (2) for any set x and object y, if $P(x)$ and $P(y)$ then we have $P(x \cup \{y\})$, then we have (3) for any z

we have $P(z)$. It should also be noted that it is an immediate consequence of Axiom D that every object is a set.

Axioms A,B,C should be recognizable as consequences of Axioms I, II, and V. Axiom D should remind one of the principle of mathematical induction.

Though we do not have a formal definition of the notion of a finite set, you should notice that if the propositional function $P(x)$ is "x is a finite set", the conditions of axiom D apply, so we should expect that everything is actually a finite set in the system described by axioms $A - D$.

Now the punchline is that axioms I-VI (but not axiom VII) are all consequences of axioms A-D.

I: is easy: it is the same as axiom A.

II: axiom B says that 0 and $\{a\}$ exist; axioms B and C give us $\{a\} + \{b\} = \{a, b\}$.

III: Let $Q(x)$ be a propositional function. We show using axiom D that the predicate $P(x)$ asserting that $\{y \in x : Q(y)\}$ exists holds of every x.

$P(0)$ is true because $\{y \in 0 : Q(y)\} = 0$.

$P(\{a\})$ is true (whether or not $P(a)$ is true) because $\{y \in \{a\} : Q(y)\}$ is either 0 or $\{a\}$ depending on whether $Q(a)$ is true, and both of these sets exist.

Suppose $P(A)$ is true. Then whether or not $P(b)$ is true $\{y \in A + \{b\} : Q(y)\}$ exists because it is either $\{y \in A : Q(y)\}$ (if $\neg Q(b)$) which exists by hypothesis, or it is $\{y \in A : Q(y)\} \cup \{b\}$, if $Q(b)$ is true, which exists by hypothesis and axiom C.

Thus Axiom III (for the specific propositional function $Q(x)$, but we made no special assumptions about it, so the argument is completely general) holds.

V: Fix a set a. $a \cup 0 = a$ exists. $a \cup \{b\}$ exists by axiom C. $a \cup (b \cup \{x\}) = (a \cup b) \cup \{x\}$ exists by axiom C if $a \cup b$ exists. So by axiom D $a \cup x$ exists for any set x.

We use this to prove the existence of general unions of sets: $\bigcup 0 = 0$. $\bigcup \{a\} = a$. $\bigcup (A \cup \{b\}) = (\bigcup A) \cup b$, and we have shown that binary unions exist just above, so by Axiom D $\bigcup x$ exists for every x.

IV: We show the existence of the set $\{y \cup \{x\} : y \in A\}$ for any set A and object x. $\{y \cup \{x\} : y \in 0\} = 0$. $\{y \cup \{x\} : y \in \{u\}\} = \{u \cup \{x\}\}$ which exists by application of Axiom C followed by Axiom B. Suppose that $\{y \cup \{x\} : y \in A\}$ exists. Then $\{y \cup \{x\} : y \in A \cup \{b\}\}$ is the union of $\{y \cup \{x\} : y \in A\} \cup \{b \cup \{x\}\}$, which exists by hypothesis that the set exists for A and applications of axioms B and C. By application of axiom D, $\{y \cup \{x\} : y \in A\}$ exists for every A.

$\mathcal{P}(0) = \{0\}$ which exists. $\mathcal{P}(\{a\}) = \{\{a\}\} \cup \{0\}$, which exists. Suppose $\mathcal{P}(A)$ exists. $\mathcal{P}(A \cup \{x\})$ is the union of $\mathcal{P}(A)$ and the set $\{a \cup \{x\} : a \in \mathcal{P}(A)\}$, which we just showed to exist. So by Axiom D every set has a power set.

VI: We want to prove that if P is a partition, P has a choice set. If $P = 0$, 0 is a choice set for P. If $P = \{a\}$, a is a nonempty set with an element x and $\{x\}$ is a choice set for $\{a\}$. If $P = A + \{b\}$ is a partition and A has a choice set C, then b is a nonempty set with an element x and $C \cup \{x\}$ is a choice set for $A + \{b\}$. By Axiom D, every partition has a choice set.

VII: We argue that no set can have its elementwise image under the singleton operation as a proper subset. This shows that Axiom VII is false if Axioms A-D hold, because an inductive set will have its elementwise image under singleton as a proper subset. The exact property $P(x)$ which we consider is the property that no subset of x has its elementwise image under the subset operation as a proper subset.

0 has this property (its elementwise image under the singleton map is 0, which is not a proper subset of 0, and it has no other subsets).

$\{a\}$ cannot have its elementwise image under the singleton operation as a proper subset, because its only proper subset is 0, and its image under the singleton operation is the nonempty set $\{\{a\}\}$. The only other subset of $\{a\}$ is the empty set already covered.

Now suppose that $A \cup \{b\}$ has a subset B whose image under the elementwise application of the singleton operation is a proper subset of B, and that A does not have this property.

Let B be a subset of $A \cup \{b\}$ which has its elementwise image under the singleton map as a proper subset. Clearly $b \in B$, as otherwise we would have a subset of A with this property contrary to hypothesis. Clearly b is the singleton of some $c \in A$, or $B \setminus \{b\}$ would have the property under consideration. We now consider the intersection C of all sets which contain $\{b\}$ and are closed under singleton (B is such a set, so we can construct C as the intersection of all subsets of B with this property). If C does not contain b as an element, it is a subset of A which has a proper subset as its elementwise image under the singleton operation. if C does contain b as an element, then C is its own image under the singleton operation (the set of singletons of elements of C would contain $\{b\}$ and be closed under the singleton operation) so $B \setminus C$ would have its own elementwise image under the singleton map as a proper subset: $B \setminus C$ certainly includes its elementwise image under the singleton map in this case, and it must include it properly or else the singleton image of B would be all of B, contrary to hypothesis. In this argument we are able to make fancy use of Axiom III because we have shown that Axiom III follows from Axioms A-D.

So in fact all the axioms I-VI hold (and axiom VII does not) if we allow only the construction of finite sets by listing.

We do not want to restrict ourselves in this way, since we do want to consider things like the collection of natural numbers and the collection of real numbers, and their general subsets.

another incidental Holmes remark about Axiom VII: Some pathologies of Zermelo set theory from a modern standpoint would be repaired by formulating Axiom VII in a more general way: Let $F(x)$ be any object-valued function (any formal definition of an object which gives an object $F(x)$ for each object x in a well-defined way), and let a be any object. Then there is a set I_F such that $a \in I_F$ and for each $x \in I_F$ we have $F(x) \in I_F$. Zermelo's axiom is of course the instance of this with $F(x) = \{x\}$.

We could then define an F, a-inductive set as a set X such that $a \in X$ and for each $x \in X$ we have $F(x) \in X$. And we can then define I_F^0, or informally $\{F^n(a) : n \in \mathbb{N}\}$, as the intersection of all

F, a-inductive elements of $\mathcal{P}(I_F)$.

The theory with this axiom is somewhat stronger, though not nearly as strong as ZFC, and it allows construction of sets such as V_ω and the transitive closure of a general set which Zermelo set theory does not allow one to define. The axiom in this form could be called the Axiom of Iteration.

The Axiom of Replacement which is usually adjoined to Zermelo to address such concerns, and which is far stronger, can be stated in the same style: for any object-valued function $F(x)$ (meaning simply a well-defined expression which picks out an object for each x) and any set A, the set $F``A$ defined as $\{F(a) : a \in A\}$ (the set of all $F(a)$ for $a \in A$) exists. This looks innocent but is amazingly strong. It is worth noticing that if the graph of $y = F(x)$ is given as a propositional function $G(x, y)$ of two variables, $F(x)$ can be defined in terms of $G(x, y)$ as $F(x) = \bigcup\{y : G(x, y)\}$, under the further condition that $G(x, y)$ implies that y is a set. My reasons for noting this are similar to Frege's reasons for introducing his $\backslash$ operator. $\{y : G(x, y)\}$ will presumably exist, since the fact that G is a graph implies that for any given x, $G(x, y)$ is true of no more than one object.

The Axiom of Replacement introduces a whole new very strong idea; the Axiom of Iteration simply generalizes what Zermelo was doing with his axiom of infinity. The use of Frege-style object valued functions as well as propositional functions in the background pleases me.

14. In section 14 we define the actual set of natural numbers, which we might call $\mathbb{N}$ and which Zermelo calls Z_0.

    We define an *inductive set* as a set which contains 0 and for each a contains $\{a\}$ if it contains a. So Axiom VII simply asserts the existence of an inductive set.

    That a set is inductive is a definite property.

    There is not necessarily a set of all inductive sets, but by Axiom III there is a set I_Z of all inductive elements of $\mathcal{P}(Z)$. Note that $Z \in I_Z$, so I_Z is nonempty. Define Z_0, or $\mathbb{N}$, as $\bigcap I_Z$. It should then be observed that for any inductive set Z', whether it is a subset of Z or not, $Z \cap Z'$

is inductive and belongs to I_Z, and so $Z_0 \subseteq Z' \cap Z \subseteq Z$, from which we see that Z_0 is the intersection of *all* inductive sets.

We informally define a function from natural numbers (whatever they may be) to their representatives in Zermelo's theory: define #0 as the set 0 and recursively define $\#(n+1)$ as $\{\#n\}$. If being an image under # is a definite predicate, it is clearly inductive. Further, by ordinary mathematical induction, it has no inductive proper subsets, so it actually is Z_0. The final step is to suggest that the answer to what the natural numbers are might be that they *are* the Zermelo natural numbers (the elements of Z_0).

remark: This ends the axiomatics part of the paper. What follows is the theory of equivalence, basically foundations of the theory of infinite cardinal number.

15. Two disjoint sets M and N are said to be *immediately equivalent*, $M \sim N$, if there is a subset Φ of their product MN such that each element of $M + N = M \cup N$ occurs as an element in one and only one element of Φ.

    Such an element Φ is called a mappinng of M onto N (notice that this is a symmetric notion, unlike our usual treatment).

    If $\{m, n\} \in \Phi$ we say that each element is mapped to or corresponds to the other.

    It should be clear that existence of Φ gives us a one to one correspondence between elements of M and elements of N: each element m of M belongs to exactly one element of Φ, which itself contains a unique element of N, which we might by an abuse of notation write $\Phi`m$. Similarly, each element of $N$ is associated with a unique $\Phi`n$. This notation is a new idea, I will see if it is useful in my exposition. It is useful to note that $\Phi`\Phi`x = x$ for each $x \in M + N$.

    $\Phi`x = \bigcup\{y \in \bigcup \Phi : y \neq x \wedge \{x, y\} \in \Phi\}$.

    Now if we have an informal one to one correspondence $f : M \to N$, we can define Φ as $\{\{x, f(x\} \in MN : x \in M\}$: this requires that f be definite in a suitable sense.

16. He begins this section by observing that it is definite for any disjoint sets M and N and subset Φ of $M + N$ and $x \in M + N$ whether the set

of elements of Φ that contain x has one element. Thus it is also evident whether all elements of $M + N$ have this property, that is, whether Φ is a mapping from M to N. Thus by Axiom III we can define a subset of $\mathcal{P}(M + N)$ containing exactly the mappings from M to N, and it is definite whether this set is 0 or not. Thus it is definite for sets M, N whether $M \sim N$. This section is entirely about definiteness and should be read closely for Zermelo's intentions. What interests me is whether he ever considers a quantifier over all of $\mathbf{B}$, or over some proper class, to be definite.

17. If Φ is a mapping from M onto N then each subset M_1 of M is mapped to a subset N_1 of N by a subset Φ_1 of Φ. The set Φ_1 can be defined as $\{x \in \Phi : x \cap M_1 \neq 0\}$. The set N_1 can be defined as $N \cap \bigcup \Phi$. Now each element of M_1 occurs in only one element of M_1, because it occurs in at least one by definition of Φ_1, and if it appeared in more than one it would also appear in more than one element of Φ. Similarly, an element of N_1 appears in at least one element of Φ_1 by definition of $\Phi_!$ and in at most one because it occurs in at most one element of Φ.

    On our own hook, we introduce the notation $\Phi\lceil M_1$ for Φ_1 and $\Phi``M_1$ for N_1.

18. If M is disjoint from N and $M \sim R$ and $N \sim R$ then $M \sim N$. Similarly, if $M \sim R$, $R \sim R'$ and $R' \sim N$, with M and N disjoint, then $M \sim N$.

    If Φ is a mapping from M onto R and Ψ is a mapping from R onto N, we define $\Phi \circ \Psi$ as $\{\{x, \Psi`(\Phi`(x))\} \in MN : x \in M\}$. Here my notation helps. It should be clear that $\{\{x, \Phi`(\Psi`(x))\} \in MN : x \in N\}$ is exactly the same set, and a mapping from M onto N.

    We define the composition of three mappings in the same way and draw the same conclusions.

19. **Theorem:** If M and N are any two sets, there is always a set M' which is equivalent to M and disjoint from N.

    Proof: Let $S = \bigcup(M + N)$ and let $r = \{u \in M + S : u \notin u\}$, for which we know by the theorem of section 10 that $r \notin M + S$. Then the sets M and $R = \{r\}$ are disjoint, and the product $M' = MR$ has the property required to witness the theorem.

Suppose that some $\{m, r\} \in MR$ belongs to N. Then $r \in \bigcup N \subseteq M + S$, because $\bigcup N \subseteq \bigcup(M + N) = S$. And this is impossible because $r \notin M + S$. This verifies that M' is disjoint from N.

That $M \sim M' = MR$ is witnessed by the mapping

$$\{\{m, \{m, r\}\} \in MM' : m \in M\}.$$

We do need to note that M and M' are disjoint, because if some $\{m, r\} \in M$ we would have $r \in \bigcup M \subseteq M + S$, because $\bigcup M \subseteq \bigcup(M + N) = S$.

Corollary: He notes specifically here that it is thus shown that no set T can contain all sets equivalent to a nonempty set M, because for any set T there is a set equivalent to M and disjoint from $\bigcup T$ and so (because nonempty) not an element of T. This makes the point that Frege's (or Cantor's) definition of cardinal number is not compatible with this theory of sets. It is historically significant that he makes this point.

An obvious observation: It is fascinating that Zermelo here makes explicit practical use of the Russell paradox diagonalization.

A further Holmes observation: This implies further, though Zermelo does not draw this conclusion, that we have a general representation of all functions from M to N. Let $F(x)$ be any function such that for any $x \in M$ we have $F(x) \in N$. Let r be defined as above. Then we can encode the function F by the set $\{\{\{m, r\}, F(m)\} \in M'N : x \in M\}$. For any $x \in M$, there is just one element of f such that $\{\{m, r\}, n\} \in N$ (there is no converse result for elements of N). We can define $f`_{M \to N}(x)$ as $\bigcup\{y \in N : x \in M \wedge \{\{x, r\}, y\} \in N\}$. Notice that we arrange for this to be 0 if $x \notin M$, and that the apparent extra parameter r is actually exactly determined by M and N.

Implicit in this is a definition of a peculiar Cartesian product $M \times N = \{\{\{x, r\}, y\} \in M'N : x \in M \wedge y \in N\}$ and a definition of set representations of functions in basically the modern way. The definition of the ordered pair suggested here is weird because it depends on the intended domain and codomain of the function, but this might even be thought to have virtues.