3. Axiomatic set theory. Axioms of Zermelo and Fraenkel

The discovery of the antinomies made it clear that a revision of the principles of set theory was necessary. The attempt to improve set theory which is best known among mathematicians is the axiomatic theory first set forth by Zermelo. I shall expose his theory in a somewhat more precise form, replacing his vague notion "definite Aussage" (= definite statement) by the notion proposition or propositional function in the first order predicate calculus. We assume that we are dealing with a domain D of objects together with the membership relation ϵ, so that all propositions are built up from atomic propositions of the form $x \epsilon y$ by use of the logical connectives &, v, -, $\rightarrow$(and, or, not, if - when) and the quantifiers (x), (Ex) (for all x, for some x). Then the following axioms are assumed valid. I write them both in logical symbols and in ordinary language.

1. Axiom of extensionality.
   If x and y have just the same elements, then x = y. In symbols

$$(z)(z \epsilon x \leftrightarrow z \epsilon y) \rightarrow (x = y)$$

   Here x = y has the usual meaning, so that

$$(x = y) \rightarrow (U(x) \leftrightarrow U(y)),$$

   where U is an arbitrary predicate. Hence we also have

$$(x = y) \rightarrow (z)(x \epsilon z \leftrightarrow y \epsilon z)$$

2. Axiom of the small sets.
   a) There exists a set without elements denoted by the symbol 0. Because of 1. there can be only one such set.

$$(E x)(y)(y \bar{\epsilon} x).$$

   b) For every object m in D there exits a set $\{m\}$ containing m, but only m, as element,

$$(x)(Ey)(x \epsilon y \ \& \ (z)(z \epsilon y \rightarrow (z = x)))$$

   c) For all m and n in D there exists a set $\{m, n\}$ containing m and n, but only these, as elements.

$$(x)(y)(Ez)(x \epsilon z \ \& \ y \epsilon z \ \& \ (u)(u \epsilon z \rightarrow (u = x) \vee (u = y)))\ .$$

   Of course b) might be omitted because it follows from c) by putting n = m.

3. Axiom of separation.
   Let C(x) be a propositional function with x as the only free variable, and m an arbitrary set. Then there exists a set consisting of all elements x of m having the property C(x).

$$(x)(Ey)(z)(z \epsilon y \leftrightarrow C(z) \ \& \ z \epsilon x)$$

4. Axiom of the power set.
   For every set m there exists a set Um whose elements are just all subsets of m.

$$(x)(Ey)(z)(z\epsilon y \leftrightarrow (u)(u\epsilon z \rightarrow u\epsilon x))$$

5. Axiom of the union.
   For every set m there exists a set Sm whose elements are just all elements of the elements of m.

$$(x)(Ey)(z)(z\epsilon y \leftrightarrow (Eu)(z\epsilon u \ \& \ u\epsilon x))$$

6. The axiom of choice.
   Let T be a set whose elements are mutually disjoint sets A,B,C,... $\neq 0$. Then there exists a set M having just one element in common with each of the sets A,B,C,...

$$(x)((y)(z)(y\epsilon x \ \& \ z\epsilon x \ \& \ y \neq z \rightarrow (u)(u\bar{\epsilon}x \ v \ u\bar{\epsilon}y))\rightarrow (Ev)((w)(w\epsilon x \rightarrow$$

$$(Et)(t\epsilon v \ \& \ t\epsilon w \ \& \ (s)(s\epsilon v \ \& \ s\epsilon w \rightarrow s = t)))).$$

These are the most general axioms set up by Zermelo (1908). Most of the general theorems of set theory are proved by the aid of these axioms. However, in order to ensure the existence of infinite sets Zermelo added:

7. The axiom of infinity.
   There exists a set U such that $0\epsilon U$ and whenever $x\epsilon U$, $\{x\}$ is ϵU as well.

$$(Ex)(0\epsilon x \ \& \ (y)(y\epsilon x \rightarrow \{y\} \epsilon x)).$$

Later Fraenkel introduced a further axiom which is more powerful with regard to the proof of the existence of large transfinite cardinals, namely the following.

8. Let the binary relation F(x,y) (= propositional function of two free variables x,y and any number of bound variables derived from the membership relation by the means of the predicate calculus) be such that $(x)(y)(z)(F(x,z) \ \& \ F(y,z)\rightarrow (y = x))$. Then to every set m there exists a set n such that $x\epsilon n \leftrightarrow (Ey)(y\epsilon m \ \& \ F(x,y))$. Or written more completely:

$$(u)(v)(w)(F(u,w) \ \& \ F(v,w) \rightarrow (u = v)) \rightarrow (x)(Ey)(z)(z\epsilon y \leftrightarrow (Eu)(u\epsilon x \ \& \ F(z,u)).$$

The following development of the Zermelo-Fraenkel set theory is carried out in such a way that it could be formalized in the predicate calculus. Such a procedure would however be very cumbersome if it were performed in all details. Therefore I have chosen an exposition that is somewhat more informal and more like the ordinary mathematical procedures.

Theorem 1. $(x)(Ey)(y\bar{\epsilon}x)$.

That means that to each set M we may find an object a such that $a\bar{\epsilon}M$. Therefore the total domain D is not a set.

Proof: According to the axiom of separation, the $x\epsilon M$ for which $x\bar{\epsilon}x$ is true, constitute the diverse elements of a set N. Then $N\bar{\epsilon}M$. Otherwise $N\epsilon N$ would imply $N\bar{\epsilon}N$ and inversely.

Theorem 2. *To each M and N there is an M' such that M' ~ M and M'* $\cap$ *(M* $\cup$ *N) = O.*

Proof: Let a be $\bar{\epsilon}$ S(M ∪ N).

The pairs $\{a,m\}$, where m runs through M, constitute a set M' obviously $\sim$ M because the pairs $(m,\{a,m\})$ furnish a one-to-one correspondence between M and M'. Indeed if $\{a,m_1\} \neq \{a,m_2\}$, then $m_1 \neq m_2$, and if $m_1 \neq m_2$, then $\{a,m_1\} \neq \{a,m_2\}$, because else we must have $m_1 = m_2$ or $m_1 = a$ & $m_2 = a$, whence again $m_1 = m_2$. Now M' is disjoint to M ∪ N, because otherwise we would have an element m of M such that $\{a,m\} \epsilon M \cup N$, whence $a \epsilon S(M \cup N)$, contrary to supposition.

Theorem 3. *Let T be a set of sets A,B,C,.... $\neq 0$ Then there exists a set T' of sets A',B',C',... together with a one-to-one correspondence between T and T' such that the unions ST and ST' are disjoint while A',B',C',... are mutually disjoint and resp. $\sim$ A,B,C,....*

Proof: According to the previous theorem a set P exists which is disjoint to T ∪ ST, while P $\sim$ T, which means that we have a one-to-one mapping $f(X) = X''$ such that X'' runs through P when X runs through T. For every $X \epsilon T$ the pairs

$$\{f(X),\ x\},$$

where x runs through X, constitute a set F(X). The function F has an inverse. Indeed, as often as $X_1 \neq X_2$, $F(X_1)$ and $F(X_2)$ will be disjoint, because $f(X_1) \neq f(X_2)$, and if we compare two elements from $F(X_1)$ and $F(X_2)$, namely

$$\{f(X_1),\ x_1\} \text{ and } \{f(X_2),\ x_2\},$$

we cannot have $f(X_1) = x_2$, because X_2 and P are disjoint. Therefore F and its inverse F' give a one-to-one correspondence between T and T' when T' is the set of all $F(X) = X'$, X running through T. For every $X \epsilon T$, the pair

$$\{f(X),\ x\}\ \epsilon X'$$

will correspond uniquely to $x \epsilon X$. If this pair is called $g_X(x)$, then g_X and its inverse yields a mapping between X and X'. In this way we have obtained a simultaneous mapping of the elements of X and those of X' for all X.

Thus the theorem is proved. However we may add the following remark: The function g is such that if $x \epsilon X$ then $x' = g_X(x)$ is $\epsilon X'$ and $x \epsilon x'$.

We have: To every $x \epsilon X$ the $x' = g_X(x)$ is the element of X' such that $x \epsilon x'$, and inversely if $x' \epsilon X'$ is given, the $x \epsilon X$ such that $g_X(x) = x'$ is the element of X which is $\epsilon x'$. The simultaneous mapping of the elements x in the diverse X onto the elements x' of the diverse X' is therefore here constructed so that $x \epsilon x'$ when x and x' correspond.

Now according to the axiom of choice there exists a set W having just one element in common with every set X'. If this element is denoted by w(X'), being a function of X' (this function is the set of pairs (X', x') where $x' = W \cap X'$), then we have

$$W \cap X' = \{w(X')\}$$

and $g_X^{-1}(w(X')) \epsilon X$, i.e. $g_X^{-1}(w(F(X))) \epsilon X$. Thus we have found a function, namely $g_X^{-1} wF$, of the elements of T which has as its value for each X an element of X. This is the general principle of choice.

Even without the axiom of choice we can introduce addition and multiplication of cardinals although only in the case of a finite number of operands. Indeed if ϕ is a set of ordered pairs (a,a') yielding a mapping of a set A onto a set A', ψ a similar set furnishing a mapping of B onto B', $A \cap B = 0$, $A' \cap B' = 0$, then $\phi + \psi$ is a mapping of $A + B$ onto $A' + B'$. Therefore we can just as in the case of the naive set theory define the sum of the cardinals of two disjoint sets as the cardinal of the sum. Similar remarks are valid for multiplication.

If we take the more general case, however, of addition, where the number of cardinal numbers to be added together is infinite; then the definition of addition is only possible when the axiom of choice is presupposed. If T is a set of mutually disjoint sets A,B,C,...., T' a set of disjoint sets A',B',C'..., while F is a mapping of T onto T' consisting of the pairs (A,A'), (B,B'),...., then if $A \sim A'$, $B \sim B'$,.... we can prove by the axiom of choice that the union ST is $\sim$ ST'. Indeed according to supposition there is a set ϕ_A of mappings of A onto A', a set ϕ_B of mappings of B onto B',... Then according to the axiom of choice there exists a set consisting of one element φ_A from ϕ_A, φ_B from ϕ_B,.... and the union of these is then a mapping ϕ of ST on ST'. Without the axiom of choice we can only formulate the following theorem: Let T and T' be as mentioned above, and let us assume that a set of mappings is given consisting of just one mapping of A onto A', one of B onto B', etc., for all elements X resp. X' of T resp. T'; then ST $\sim$ ST'.

There is on the other hand one important theorem concerning the comparison of cardinals which can be proved without the axiom of choice, namely the Bernstein Theorem.

Theorem 4. *Let M be $\sim M'$, $M' \subset M_1 \subset M$. Then $M \sim M_1$.*

Remark: I use for every subset A of M the notation A' for the image of A by the same mapping as of M onto M'.

Proof: We put

$$M_1 = Q + M', \text{ or in other words } Q = M_1 - M'.$$

Let T be the set of subsets A of M which have the properties

1) $Q \subseteq A$ 2) $A' \subseteq A$.

T is not empty because at least $M \epsilon T$. Then let A_0 be the intersection of all elements of T. I denote this also by DT. Obviously A_0 has still the properties 1) and 2), i.e., $A_0 \epsilon T$ or

3) $Q \subseteq A_0$ 4) $A_0' \subseteq A_0$.

3) and 4) furnish 5) $Q \cup A_0' \subseteq A_0$

whence $$(Q \cup A_0')' \subseteq A_0'$$

whence a fortiori

$$6) \quad (Q \cup A_0')' \subseteq Q \cup A_0' .$$

From 6) it follows that

$$Q \cup A_0' \epsilon T,$$

whence

$$7)\quad A_0 \subseteq Q \cup A_0',$$

5) and 7) yield

$$A_0 = Q + A_0',$$

noticing that $Q \cap A_0' \subseteq Q \cap M' = 0$. Now we have

$$M_1 = Q + M' = Q + A_0' + (M' - A_0') = A_0 + (M' - A_0'),$$

whence, A_0 being $\sim A_0'$,

$$M_1 \sim A_0' + (M' - A_0') = M'$$

which is the theorem.

An immediate consequence is that if

$$M \sim N_1 \subset N \quad\text{and}\quad N \sim M_1 \subset M,$$

then

$$M \sim N.$$

Indeed it follows from $N_1 \subset N$ and $N \sim N_1$ that $N_1 \sim M_2$, where M_2 is a certain subset of M_1, so that since $M \sim N_1$

$$M \sim M_2 \subset M_1 \subset M,$$

whence after the previous theorem

$$M \sim M_1 \sim N.$$

Corollary: If $M \sim M' \subset M$, $\mathfrak{m} = \overline{\overline{M}}$, then

$$\mathfrak{m} + 1 = \mathfrak{m}.$$

It may be remarked that we have not used the axiom of choice in the proof of this theorem. As an example of another simple theorem of a certain interest, provable as well without the axiom of choice, I will mention Cantor's theorem and the very simple one below concerning the case $\mathfrak{m}$ and $\mathfrak{n} \geqq 2$.

Theorem 5. (Cantor's theorem). *For every set M we have* $\overline{\overline{M}} < \overline{\overline{UM}}$.

Proof: In the first place the pairs $(m, \{m\})$ yield a mapping of M on a subset of UM, namely the subset consisting of all sets $\{m\}$ where $m \in M$. In the second, no mapping f of UM into M can exist. Indeed, let us assume the existence of such a mapping f and let N be the set of all f(X) for subsets X of M for which $f(X)\bar{\epsilon}X$. Then we should have

$$(X)(X \subseteq M \rightarrow (f(X) \epsilon N \leftrightarrow f(X)\bar{\epsilon}X).$$

Putting in particular N into this formula instead of X, we obtain, since $N \subseteq M$,

$$f(N) \epsilon N \leftrightarrow f(N)\bar{\epsilon}N$$

which is absurd.

Using the cardinal number notation this theorem may be written

$$2^{\mathfrak{m}} > \mathfrak{m},$$

because it is seen that the cardinal number of UM must be $2^{\mathfrak{m}}$ when $\mathfrak{m}$ denotes the cardinal number of M. This is perhaps seen most convincingly in the following way. Let M' be $\sim M$ and $M \cap M' = 0$, f being a mapping of M onto

M'. We know that such M' and f exist. For every $m \in M$ I write $f(m) = m'$. Then we can get a one-to-one correspondence between UM and the product of all the pairs $\{m,m'\}$. Let N be $\subseteq M$. Then as often as $m \in M$ is also $\in N$ we let the corresponding element of the product contain m as element, otherwise it contains just m'. Since the set of all pairs $\{m,m'\}$ evidently has the same cardinal number $\mathfrak{m}$ as M the product must be of cardinality $2^{\mathfrak{m}}$.

A consequence of this theorem is that a set of sets representing all cardinals does not exist. Indeed, if T is such a set, then $\overline{\overline{ST}} \geqq \overline{\overline{X}}$, X an arbitrary element of T, and Cantor's theorem says that $\overline{\overline{UST}} > \overline{\overline{ST}}$. Hence $\overline{\overline{UST}} > \overline{\overline{X}}$ for all $X \in T$.

It may be suitably mentioned here, that the sum of the cardinals belonging to a set of sets with no greatest cardinal has already a cardinal > all cardinals in the set.

It is often asserted that the following theorem, also due to Bernstein, can be proved without the axiom of choice. However, the usual proof, at least, does not fulfill this requirement, so that I think it is a mistake. The theorem is, when $\mathfrak{m}$ and $\mathfrak{n}$ denote cardinals:

Theorem 6. *If* $\mathfrak{m} + \mathfrak{n} = \mathfrak{m}\mathfrak{n}$, *then* $\mathfrak{m}$ *and* $\mathfrak{n}$ *are comparable.*

What is meant is that either $\mathfrak{m} \leqq \mathfrak{n}$ or $\mathfrak{n} \leqq \mathfrak{m}$.

Proof: The supposition $\mathfrak{m} + \mathfrak{n} = \mathfrak{m}\mathfrak{n}$ means that we are given two disjoint sets M and N together with a mapping of M + N onto $M \times N$. This means again that the set of all pairs (m,n), $m \in M$, $n \in N$, is divided into two disjoint parts A and B where A is mapped onto M, B onto N. Now, if there is a particular m' such that all (m',n), n running through N, are $\in A$, then N is $\sim$ a subset of A, whence N $\sim$ a subset of M. If no such m' exists, then for each $m \in M$ there is at least one n such that $(m,n) \in B$. Then one says, it is evident that B contains a subset which is $\sim M$, whence M $\sim$ a subset of N.

Theorem 7. *Let the cardinals* $\mathfrak{m}$ *and* $\mathfrak{n}$ *be* $\geqq 2$. *Then* $\mathfrak{m} + \mathfrak{n} \leqq \mathfrak{m}\mathfrak{n}$.

Proof: We have two sets M and N with at least two elements and we can assume M and N disjoint. Let $m_1 \neq m_2$ be $\in M$, $n_1 \neq n_2 \in N$. Let P be the set of all $\{m_1, n\}$, n running through N. Then P is $\sim$ N. Further let Q be the set of all $\{m,n_1\}$, m running through M - $\{m_1\}$, besides the pair $\{m_2,n_2\}$. It is evident that Q is $\sim$ M. Further P and Q are disjoint. Thus P + Q is a subset of $M \cdot N$ which is $\sim$ M + N, which proves the theorem.

It is seen that the hypothesis of the theorem can not be weakened. Indeed if it is only supposed that one of the two cardinals is $\geqq 2$, the other, say n, being = 1, the theorem is not valid for finite $\mathfrak{m}$.

The theorem can be generalized. Let T be a set of at least 2 elements, each element of T containing at least two elements, the elements of T being mutually disjoint. Using the axiom of choice we may assume that we have chosen two elements of each $X \in T$. Let A,B,C,.... be the elements of T and let $a_1, a_2, b_1, b_2, \ldots$ be the chosen elements from A,B,C,.... Then the product PT contains subsets

$$A_1 \qquad B_1 \qquad C_1 \qquad . \quad . \quad .$$

consisting of the elements

$$a_r, b_1, c_1, \ldots\ldots \qquad a_1, b_s, c_1, \ldots\ldots \qquad a_1, b_1, c_t, \ldots\ldots$$
$$r \neq 1 \qquad\qquad s \neq 1 \qquad\qquad t \neq 1$$

and $\quad a_1, b_2, c_2, \ldots\ldots \qquad a_2, b_1, c_2, \ldots\ldots \qquad a_2, b_2, c_1, \ldots\ldots$

and it is evident that $A_1 \sim A$, $B_1 \sim B$ .... This means that PT contains a subset ST so that $\overline{\overline{ST}} \leqq \overline{\overline{PT}}$.

Theorem 8. *If* $0 < \mathfrak{m} \leqq \mathfrak{n}$, *then every set of cardinality* $\mathfrak{n}$ *can be divided into a set T of cardinal* $\mathfrak{m}$ *of non-void mutually disjoint subsets.*

Proof: Let $M \subset N$, $\overline{\overline{M}} = \mathfrak{m}$, $\overline{\overline{N}} = \mathfrak{n}$ and $m \in M$. For each $x \in M$ a subset N_x of N is defined thus: If $x \neq m$, then $N_x = \{x\}$, while in the case $x = m$, $N_m = (N-M) + \{m\}$. It is evident that the N_x, x running through M, are all mutually disjoint and their union (sum) is N.

The inverse of this would be, that if a set N is the union of a set T of non-void mutually disjoint sets, then $\overline{\overline{T}} \leqq \overline{\overline{N}}$. However, without axiom of choice this can only be proved if T is finite. Indeed, in order to prove this assertion one has to find a subset of N which is $\sim$ T. This is possible if we can choose one element a from each element A of T; then the pairs (a,A) yield a mapping of T on a subset of N. Otherwise we have no means of proof. On the other hand we may prove the following theorem. Let N be the sum of mutually disjoint and non-void sets N_x, $x \in M$, so that to each $x \in M$ corresponds just this single N_x. Then M is $\sim$ a subset of the power set of N, so that $\mathfrak{m} = \overline{\overline{M}} \leqq 2^{\mathfrak{n}}$, $\mathfrak{n} = \overline{\overline{N}}$. To every subset $X \subset M$ we let correspond the subset N_X, namely the sum of all N_x, x running through X, which is $\subset N$. For different X these corresponding N_X are different; therefore $2^{\mathfrak{m}} \leqq 2^{\mathfrak{n}}$. If we had $\mathfrak{m} = 2^{\mathfrak{n}}$, then $2^{\mathfrak{m}} \leqq \mathfrak{m}$ which is not the case, by Cantor's theorem. Thus $\mathfrak{m} < 2^{\mathfrak{n}}$.

Theorem 9. (Zermelo). *Let T be mapped on T' in such a manner that as often $M \in T$ corresponds to $M' \in T'$, $\overline{\overline{M}} < \overline{\overline{M'}}$. Then $\overline{\overline{ST}} < \overline{\overline{PT'}}$.*

Proof: We may assume that the elements A,B,C,... of T are $\neq 0$. Then $\overline{\overline{A'}}, \overline{\overline{B'}}$ ... are all $\geqq 2$. By theorem 7 we then know that $\overline{\overline{ST'}} \leqq \overline{\overline{PT'}}$. Further it is clear that $\overline{\overline{ST}} \leqq \overline{\overline{ST'}}$. Thus $\overline{\overline{ST}} \leqq \overline{\overline{PT'}}$ and it suffices to prove that PT' cannot be mapped on a subset Σ of ST. Let us assume that such a mapping were possible. The subset Σ of ST can be written as $A_0 + B_0 + C_0 + \ldots$, where A_0 is the intersection of Σ and A, B_0 that of Σ and B,.... The elements of PT are of the form $\{a',b',c',\ldots\}$, where $a' \in A'$, $b' \in B'$,.... Let us take into account those which correspond to the elements of A_0. If a' varies, the corresponding $a \in A_0$ varies. Therefore the a' occurring in the elements $\{a',b',c',\ldots\}$ which are mapped on the elements of A_0 can only constitute a proper subset A_1 of A', because else A' would have to be $\sim A_0$ which contradicts the assumption $\overline{\overline{A}} < \overline{\overline{A'}}$. Similarly the b' occurring in the elements $\{a',b',c',\ldots\}$ which are mapped on the elements of B_0 must constitute a proper subset B_1 of B', and so on. Now PT also contains, according to the axiom of choice, an element,

$$\{a_0, b_0, c_0, \dots\},$$

where $a_0 \epsilon A' - A_1$, $b_0 \epsilon B' - B_1$, ... However this element cannot correspond to any element of ST. Indeed it cannot be mapped on an element of A_0, for example, because if it could, a_0 would have to be one of the elements of A_1.

4. The well-ordering theorem

After all this I shall now prove, by use of the choice principle, that every set can be well-ordered. First I shall give another version of the notion "well-ordered", different from the usual one.

We may say that a set M is well-ordered, if there is a function R, having M as domain of the argument values and UM as domain of the function values, such that if $N \supset 0$ is arbitrary and ϵ UM, there is a unique $n \epsilon N$ such that $N \subseteq R(n)$. I have to show that this definition is equivalent to the ordinary one. If M is well-ordered in the ordinary sense, then every non-void subset N has a unique first element. Then it is clear that if R(n), $n \epsilon M$, means the set of all $x \epsilon M$ such that $n \leqq x$, the other definition is fulfilled by this R. Let us, on the other hand, assume that we have a function R of the said kind. Letting N be $\{a\}$, one sees that always $a \epsilon R(a)$. Let N be $\{a,b\}$, $a \neq b$. Then either a or b is such that $N \subseteq R(a)$ resp. R(b). If $N \subseteq R(a)$, then we put $a < b$. Since then N is not $\subseteq R(b)$, we have $a \bar{\epsilon} R(b)$. Now let $b < c$ in the same sense that is, $c \epsilon R(b)$, $b \bar{\epsilon} R(c)$. Then it is easy to see that $a < c$. Indeed we shall have $\{a,b,c\} \subseteq$ either R(a) or R(b) or R(c), but $b \bar{\epsilon} R(c)$, $a \bar{\epsilon} R(b)$. Hence $\{a,b,c\} \subseteq R(a)$ so that $\{a,c\} \subseteq R(a)$, i.e. $a < c$. Thus the defined relation $<$ is linear ordering. Now let N be an arbitrary subset of M and n be the element of N such that $N \subseteq R(n)$. Then if $m \epsilon N$, $m \neq n$, we have $m \epsilon R(n)$, which means that $n < m$. Therefore the linear ordering is a well-ordering.

Theorem 10. *Let a function ϕ be given such that $\phi(A)$, for every A such that $O \subset A \subseteq M$, denotes an element of A. Then UM possesses a subset $\mathfrak{M}$ such that to every $N \subseteq M$ and $\supset O$ there is one and only one element N_0 of $\mathfrak{M}$ such that $N \subseteq N_0$ and $\phi(N_0) \epsilon N$.*

Proof: I write generally $A' = A - \{\phi(A)\}$. I shall consider the sets $P \subseteq$ UM which, like UM, possess the following properties

1) $M \epsilon P$

2) $A \epsilon P \rightarrow A' \epsilon P$ for all $A \subseteq M$

3) $T \quad P \rightarrow DT \epsilon P$.

These sets P constitute a subset $\mathfrak{T}$ of UUM. They are called Θ-chains by Zermelo. I shall show that the intersection $D\mathfrak{T}$ of all elements of $\mathfrak{T}$ is again a Θ-chain, that is, $D\mathfrak{T} \epsilon \mathfrak{T}$. It is seen at once that $D\mathfrak{T}$ possesses the properties 1) and 2). Now let $T \subseteq D\mathfrak{T}$. Then, if $P \epsilon \mathfrak{T}$, we have $T \subseteq P$, and since 3) is valid for P, also $DT \epsilon P$. Since this is true for all P, we have $DT \epsilon D\mathfrak{T}$ as asserted. Thus I have proved that $D\mathfrak{T} \epsilon \mathfrak{T}$.