A Method for Eliminating Transfinite Numbers from Mathematical Arguments

by Casimir (Kazimierz) Kuratowski, translated by Paul Taylor

1922 and 2024

The theory of ordinals[1] may be considered from two points of view. As a significant generalisation of arithmetic, it is of deep philosophical meaning and presents by itself one of the most beautiful and interesting topics in modern mathematics. On the other hand, its applications have contributed in many ways to the advancement of different areas of our subject. Indeed, it was for the benefit of these applications that Cantor originally developed his theory [Can15, Can32].

In this work we treat this theory solely from the point of view of such applications. We show that theorems of a certain general form can actually be proved *without* making use of these numbers. This includes results from many well known and important disciplines where the *proofs* have employed ordinals, even though the *statements* make no mention of them.

The goal of eliminating ordinals from arguments that belong solely to applications is not a new one: it has been done by Lebesgue [Leb05], Lindelöf for the Cantor–Bendixson Theorem [Lin05], Sierpiński [Sie20], the Youngs for the class of derivatives of any order of a given set of points [YY06, appx] and Zermelo for the Well-Ordering Theorem [Zer04, Zer08a], but using different methods in each particular case.

However, the ways in which ordinals are employed have actually been *uniform*, so that we may give them according to a common scheme. After presenting this, *we will transform any process represented by it into another one that no longer employs ordinals.*

Our method is closely linked to Dedekind's notion of a chain ("Kette") [Ded88] that was developed by Zermelo in his second proof of the Well-Ordering Theorem [Zer08a] and by Hessenberg [Hes08].

Although the use of ordinals may sometimes give certain advantages in terms of brevity and simplicity, the existence of a general procedure for removing them from proofs of theorems that make no mention of them is important for the following reasons:

When reasoning with ordinals, one makes the implicit assumption that they *exist*, whereas minimising the system of axioms used in a proof is desirable logically and mathematically. Besides, such reduction frees the development from ideas that do not belong in it, which is desirable aesthetically.

From the point of view of Zermelo's axiomatisation of set theory [Zer08b], it seems that the method presented here allows one to prove theorems of a certain general form directly from his axioms. That is, with no need to introduce additional axioms for the existence of ordinals.

[1] Kuratowski uses "nombre transfini" throughout, but repeating this phrase seems cumbersome, so we have replaced it with "ordinal". He also refers to transfinite numbers "of the first or second kind", which we have rendered as "successor" and "limit ordinals".

The general method

Before studying the general pattern of the use of ordinals, let us consider a particular application, namely how one usually defines coherences of every order for a set of points:

Let S be Euclidean space of dimension n. For any set X of points, let X' be the derived set of X and $G(X) \equiv X \cap X'$. Let A be a given set of points and define:

- $A_0 \equiv A$;
- $A_{\alpha+1} \equiv G(A_\alpha)$;
- for any limit ordinal α, let $A_\alpha \equiv \bigcap_{\xi<\alpha} A_\xi$.

A_α is called the *coherence of order α of A*. So the class $\mathbf{A}(A)$ of all of the A_α is the class of coherences of all orders for A.

Here is the general pattern of the procedure to which we have referred:

Schema I. Let E be a given set, whose elements may be of any kind whatever. The only requirement on E is that for any subset $X \subset E$, $G(X)$ denotes a subset of E that satisfies the ["deflationary"] inclusion

$$G(X) \subset X. \tag{1}$$

Let $A \subset E$ be any given subset of E. Neither its definition nor those of the set E nor the function $G(X)$ make any reference to ordinals. Nevertheless, for all ordinals α, put

$$A_0 \equiv A, \tag{2}$$

$$A_{\alpha+1} \equiv G(A_\alpha), \tag{3}$$

and when α is a limit ordinal,

$$A_\alpha \equiv \bigcap_{\xi<\alpha} A_\xi. \tag{4}$$

Schema I′ is the symmetrical one obtained by replacing (1) by [the "inflationary" inclusion]

$$X \subset G(X) \tag{5}$$

and the equation (4) by

$$A_\alpha \equiv \bigcup_{\xi<\alpha} A_\xi. \tag{6}$$

Given a $G(X)$ function obeying Schema I, consider the families[2] $\mathbf{Z}$ such that

$$\text{the elements of } \mathbf{Z} \text{ are subsets of } E, \tag{7}$$

$$A \in \mathbf{Z}, \tag{8}$$

$$X \in \mathbf{Z} \quad \text{implies} \quad G(X) \in \mathbf{Z}, \tag{9}$$

$$\emptyset \neq \mathbf{X} \subset \mathbf{Z} \quad \text{implies} \quad \bigcap \mathbf{X} \in \mathbf{Z}. \tag{10}$$

Here we use A, B, X, *etc.* to denote [sub]sets. Families of [sub]sets are written $\mathbf{A}$, $\mathbf{B}$, ... $\bigcap \mathbf{X}$ is the intersection[3] of the subsets in $\mathbf{X}$ and $\bigcup \mathbf{X}$ their union. In particular, if the family $\mathbf{X}$ consists

[2] Kuratowski uses the word "class" for any collection of subsets of the base set E. We have translated this as "family" when it is definable in Zermelo set theory with bounded Quantification and Selection and "class" otherwise. This distinction is explained in the Translator's Note at the end.

[3] Kuratowski says "produit" and writes $\times$ and $\prod$ for what we now call intersection, $\cap$, $\bigcap$. Similarly he uses "somme", $+$ and $\sum$ for our union, $\cup$ and $\bigcup$. We have also changed his "$-$" for subset subtraction to $\setminus$.

of a sequence $X_1, X_2, \ldots X_n, \ldots$ then we also write $\bigcap_{n=1}^{\infty} X_n$ and $\bigcup_{n=1}^{\infty} X_n$ for their intersection and union.

Such families $\mathbf{Z}$ do exist; for example, the family of *all* subsets of E is one such. Among such families there is a smallest one. Indeed, their intersection is also a family $\mathbf{Z}$, since (as shown without difficulty) it satisfies conditions (7–10). Let's call this smallest family $\mathbf{M}(A)$. [It is "generated" by A, G and $\bigcap$.]

Definition. Given a set E and a function $G(X)$ such that, for any subset $X \subset E$, $G(X)$ is a subset of E and satisfies the inclusion (1) [*i.e.* $G(X) \subset X \subset E$], For any subset $A \subset E$, $\mathbf{M}(A)$ denotes the smallest family $\mathbf{Z}$ satisfying conditions (7–10).

In exactly the same way, we define $\mathbf{N}(A)$ by replacing condition (1) with (5) and (10) with

$$\emptyset \neq \mathbf{X} \subset \mathbf{Z} \quad \text{implies} \quad \bigcup \mathbf{X} \in \mathbf{Z}. \tag{11}$$

The existence of the family $\mathbf{N}(A)$ is proved in the same way.

The method of eliminating ordinals consists of *replacing the class* $\mathbf{A}(A)$ *in any process represented by the Schema I (or* I′*) with the family* $\mathbf{M}(A)$ *or* $\mathbf{N}(A)$.

As it is easy to see, the definition of the family $\mathbf{M}(A)$ and the proof its existence make no recourse to ordinals. When we consider the ways in which ordinals have been used to date, we will show by using this method and without having to invoke ordinals how to recover the results that are usually obtained with them.

In fact, in many particularly simple cases, it is possible to eliminate the ordinals directly, without the help of the general method presented here. This is, for example, the case of the following reasoning, *cf.* [Hau14, p.275].

Let $\mathbf{K}$ be an infinite well ordered descending sequence of closed sets [in Euclidean space], where it is required to demonstrate that the family $\mathbf{K}$ is countable.

Suppose that the elements of this family are arranged in a transfinite sequence, writing K_α for the general term. Let $\{R_n\}$ be the sequence of spheres with rational centres and radii. For any given K_α, let $n(K_\alpha)$ be an index chosen such that the sphere $R_{n(K_\alpha)}$ has points in common with K_α but not with the members of the family $\mathbf{K}$ earlier than K_α and $n(K_\alpha)$ is the smallest number that has this property.

We immediately see that any two different elements of $\mathbf{K}$ thus correspond to two different natural numbers, so the family $\mathbf{K}$ is countable.

By examining this argument we recognise without difficulty that we can remove the ordinals from it. For this it is enough to replace K_α with the set K, making no use of the transfinite.

If we compare this reasoning to the one used to build the class of coherences, we see the following essential difference: in the first the "chain" (in particular, the well-ordered family $\mathbf{K}$) is *given by hypothesis*, whilst that in the second it (*i.e.* the class $\mathbf{A}(A)$) has to be constructed. It is only necessary to use our *general* method of eliminating ordinals when the chain is *not given in advance* but is constructed as *part of the proof.*

In particular, the role of ordinals in

- Bernstein's results about sets without perfect parts (Leipz. Ber, 60),
- in Mahlo's about the number of "homoïes" (ibid. 63),
- in the article by Sierpiński and myself on classes ($\mathbf{L}$) [KS21] and
- in the one by Knaster and myself on connected sets [KK21]

is similar to their role in the reasoning on family $\mathbf{K}$ just discussed.

We will now show that $\mathbf{M}(A) = \mathbf{A}(A)$. Of course, if the class $\mathbf{A}(A)$ has been *defined* using ordinals, as in Schema I, this equality cannot be established without using them. However, in no individual case where there is a need to eliminate ordinals will we make use of this equation: it will never occur as a *premise*. The sole reason why we prove it here is that this identity allows us to state that our method can eliminate ordinals from *any* process of the form in Schema I (or I′).

Since the class $\mathbf{A}(A)$ is a well ordered descending class of sets, it satisfies conditions (7–10). So $\mathbf{A}(A)$ is a class $\mathbf{Z}$ and, like $\mathbf{M}(A)$ is, by definition, contained in each $\mathbf{Z}$, we have $\mathbf{M}(A) \subset \mathbf{A}(A)$. The reverse inclusion also holds. Indeed, according to (2) and (8), $A_0 \in \mathbf{M}(A)$, and if $A_\xi \in \mathbf{M}(A)$ for all $\xi < \alpha$ then $\mathbf{A}_\alpha \in \mathbf{M}(A)$. From this we deduce that $A_\alpha \in \mathbf{M}(A)$, using conditions (3) and (9) or (4) and (10), according as α is a successor or limit ordinal. Therefore $\mathbf{A}(A) = \mathbf{M}(A)$.

In a similar way we may show that the family $\mathbf{N}(A)$ is the same as $\mathbf{A}(A)$ in Schema I′.

We will now establish some general theorems about the families $\mathbf{M}(A)$ and $\mathbf{N}(A)$ that we will often use later. The proofs of these results make no use of the notion of ordinals and will be based entirely on Zermelo's axioms [Zer08b], in particular Axioms I–V.

The most important property of the class $\mathbf{A}(A)$ it that it enables transfinite induction. Therefore it is essential that such induction remain valid using the method proposed here, albeit in a modified form, for any argument using the families $\mathbf{M}(A)$ or $\mathbf{N}(A)$.

Indeed, suppose that we need to prove a given property of elements of the family $\mathbf{M}(A)$, then we must apply the following procedure. Consider the family $\mathbf{P}$ of all subsets of E that enjoy the property in question. Then show that

1. $A \in \mathbf{P}$,
2. if $X \in \mathbf{P}$ then $G(X) \in \mathbf{P}$ and
3. if $\emptyset \neq \mathbf{X} \subset \mathbf{P}$ then $\bigcap \mathbf{X} \in \mathbf{P}$.

Since the family $\mathbf{P}$ is a family $\mathbf{Z}$ that satisfies conditions (7–10), by the definition of $\mathbf{M}(A)$ we have $\mathbf{M}(A) \subset \mathbf{P}$. This says precisely that all of the elements of $\mathbf{M}(A)$ enjoy the property in question.

The process that we have just described, which we also call *induction*, will recur over and over again in what follows. Its legitimacy follows from the same definition as $\mathbf{M}(A)$.

So we have established

Theorem I. Suppose given a property $\mathbf{P}$ such that

1. the given set A obeys $\mathbf{P}$;
2. if X satisfies it then so does $G(X)$;
3. for any [non-empty] family $\mathbf{X}$ of sets that satisfy $\mathbf{P}$, so does their intersection $\bigcap \mathbf{X}$.

Then every element of $\mathbf{M}(A)$ has the property $\mathbf{P}$.

Symmetrically, we obtain **Theorem I′** by replacing $\bigcap \mathbf{X}$ with $\bigcup \mathbf{X}$ and $\mathbf{M}(A)$ by $\mathbf{N}(A)$.

[An easy consequence of Theorem I that is used later in the paper but not stated clearly is that A is the greatest member of $\mathbf{M}(A)$. This follows by letting $\mathbf{P}$ consist of those X with $X \subset A$.]

Here are some important properties of the set $\bigcap \mathbf{M}(A)$.

From (10), this set still belongs to the family $\mathbf{M}(A)$ and is its smallest member. Writing $P(A)$ for it, we therefore have $P(A) \subset M$ for any element $M \in \mathbf{M}(A)$. In particular $P(A) \subset A$, by (8). We also claim that

$$P(A) \;=\; G\big(P(A)\big). \tag{12}$$

Indeed, since $P(A)$ is an element of $\mathbf{M}(A)$, so is $G\big(P(A)\big)$, by (9). Therefore $P(A) \subset G\big(P(A)\big)$ and they are equal by inclusion (1).

Suppose now that the function $G(X)$ obeys the condition

$$X \subset Y \quad \text{implies} \quad P(X) \subset G(Y). \tag{13}$$

We will show that, under this hypothesis, $P(A)$ is the *greatest* set that, when substituted for Z, satisfies the formulae:

$$Z \;\subset\; A, \tag{14}$$

$$Z \;=\; G(Z), \tag{15}$$

So let Z be any set that satisfies (14) and (15). We must show that

$$Z \;\subset\; P(A). \tag{16}$$

Because of (15), the family $\mathbf{M}(Z)$ has just a single element Z and so

$$Z \;=\; P(Z). \tag{17}$$

Let U be a set such that

$$Z \;\subset\; U. \tag{18}$$

According to (18), (13) and (17), we have

$$Z \;=\; P(Z) \;\subset\; G(U), \tag{19}$$

so the inclusion (18) implies (19) for any U.

Using (14) and Theorem I (in which $\mathbf{P}$ denotes the property of containing Z), we deduce that every member of the family $\mathbf{M}(A)$ also contains Z, since

$$(X \in \mathbf{P}) \;\equiv\; (Z \subset X) \quad \text{implies} \quad \big(Z \subset G(X)\big) \;\equiv\; \big(G(X) \in \mathbf{P}\big).$$

In particular the inclusion (16) holds. QED

Replacing the condition (13) by the more restrictive

$$X \subset Y \quad \text{implies} \quad G(X) \subset G(Y), \tag{20}$$

we immediately deduce

Theorem II. If the function $G(X)$ [is deflationary and monotone, *i.e.* it] satisfies conditions (1) and (20), then the set $P(A)$ (*i.e.* the smallest set in the family $\mathbf{M}(A)$) is the largest subset $Z \subset A$ that is fixed by G, *i.e.* satisfies the equation (15).

Similarly, for a function $G(X)$ obeying condition (5), $S(X) \equiv \bigcup \mathbf{N}(A)$ is a Z that satisfies (15). If we suppose further that

$$X \subset Y \quad \text{implies} \quad G(X) \subset S(Y), \tag{21}$$

then $S(A)$ is the smallest Z that contains A and is a subset of E.

Since the condition (20) is more restrictive than (21), we deduce

Theorem II′. If the function $G(X)$ [is inflationary and monotone, *i.e.* it] satisfies conditions (5) and (21) then the set $S(A)$ (*i.e.* largest set in the family $\mathbf{N}(A)$) is the smallest subset $Z \subset E$ that contains A and satisfies the equation (15).

Theorem III. $\mathbf{M}(A)$ is a family of decreasing sets. That is, for any $X, Y \in \mathbf{M}(A)$,

$$\text{either} \quad X \subset Y \quad \text{or} \quad Y \subset X.$$

Proof. This is a [double] application of the induction principle that we have just derived.

Consider the elements $K \in \mathbf{M}(A)$ that satisfy the condition

$$\text{for all} \quad X, Y \in \mathbf{M}(A) \quad \text{such that} \quad K \subset X, \quad \text{either} \quad X \subset Y \quad \text{or} \quad Y \subset G(X). \tag{22}$$

Let $\mathbf{K}$ be the family of all such K. We claim that this is a family $\mathbf{Z}$ that satisfies condition (8–10).

[**Outer base case (8)**, that] $A \in \mathbf{K}$.

Consider the family $\mathbf{Y}$ consisting of those $Y \in \mathbf{M}(A)$ such that

$$Y = A \quad \text{or} \quad Y \subset G(A).$$

[Since $G(Y) \subset Y$,] this is a family $\mathbf{Z}$, satisfying conditions (8–10), so by Theorem I,

$$\mathbf{Y} \ = \ \mathbf{M}(A) \tag{23}$$

and A is a K that satisfies (22).

[**Outer successor case (9)**, that $K \in \mathbf{K}$ implies $G(K) \in \mathbf{K}$.]

For a given K consider the subfamily $\mathbf{T} \subset \mathbf{M}(A)$ whose members are the T satisfying one of the following two conditions:

$$T \ \supset \ G(K) \tag{24}$$

or

$$T \ \subset \ G\big(G(K)\big). \tag{25}$$

[**Inner base case:**] From (23), A is a super-set of all of the members of $\mathbf{M}(A)$, so we may put it for T in formula (24). Then

$$A \ \in \ \mathbf{T}. \tag{26}$$

[**Inner successor case:**] Now let T be any element of $\mathbf{T}$. We claim that $G(T) \in \mathbf{T}$.

First note that if $T \equiv K$ or $T \equiv G(K)$ then $G(T) \in \mathbf{T}$ using the formulae (1), (24) and (25). So we may suppose that

$$T \ \neq \ K, \tag{27}$$

and

$$T \ \neq \ G(K). \tag{28}$$

Two cases may occur, depending on whether T satisfies (24) or (25).

In the first case we have $T \not\subset G(K)$ by (28). So, putting $X \equiv K$ and $Y \equiv T$ in (22), we deduce that $K \subset T$. Putting $X \equiv T$ and $Y \equiv K$ in the same formula, we have $K \subset G(T)$ by (27). Hence $G(K) \subset K \subset G(T)$ and $G(T) \in \mathbf{T}$, since we may substitute $G(T)$ for T in (24).

In the second case we have $G(T) \subset T \subset GG(K)$ and, substituting $G(T)$ for T in (25), we deduce that $G(T) \in \mathbf{T}$.

We have shown that

$$T \in \mathbf{T} \quad \text{implies} \quad G(T) \in \mathbf{T}. \tag{29}$$

[**Inner limit case:**] On the other hand,

$$\emptyset \neq \mathbf{X} \subset \mathbf{T} \quad \text{implies} \quad \bigcap \mathbf{X} \in \mathbf{T}. \tag{30}$$

Indeed, if all the elements $Y \in \mathbf{X}$ satisfy (24), so does their intersection $\bigcap \mathbf{X}$. If not, at least one of them is a subset of $GG(K)$ by (25). *A fortiori*, $\bigcap \mathbf{X} \subset GG(K)$.

[**Inner conclusion:**] By Theorem I, the formulae (26), (29) and (30) entail the equation

$$\mathbf{T} \;=\; \mathbf{M}(A) \tag{31}$$

and therefore $G(K) \in \mathbf{K}$. Indeed, let $X \in \mathbf{M}(A)$ and $G(K) \subset X$. If, in addition, $K \subset X$, by(22), for any $Y \in \mathbf{M}(A)$ we have

$$X \subset Y \qquad \text{or} \qquad Y \subset G(X).$$

Otherwise we would deduce from (22) that $X \subset G(K)$ and so that $X = G(K)$; then by the formulae (31), (24) and (25), $X \subset Y$ or $Y \subset G(X)$. Hence in either case we may substitute $G(K)$ for K in (22).

[**Outer limit case (10)**, that $\emptyset \neq \mathbf{U} \subset \mathbf{K}$ implies $\bigcap \mathbf{U} \in \mathbf{K}$.]

We write $\mathbf{R}$ for the subfamily of $\mathbf{M}(A)$ consisting of the sets R with

$$\text{either} \quad \bigcap \mathbf{U} \subset R \quad \text{or} \quad R \subset G(\bigcap \mathbf{U}).$$

[**Inner base case:**] By (23), $A \in \mathbf{R}$.

[**Inner successor case:**] On the other hand, let $R \in \mathbf{R}$. If $R = \bigcap \mathbf{U}$ or $R \subset G(\bigcap \mathbf{U})$ then $G(R) \subset G(\bigcap \mathbf{U})$, whence $G(R) \in \mathbf{R}$. Suppose therefore that $R \neq \bigcap \mathbf{U}$ and $R \not\subset G(\bigcap \mathbf{U})$; then $R \supset \bigcap \mathbf{U}$ from the definition of $\mathbf{R}$. We claim that there is some element $K \in \mathbf{U}$ with $K \subset R$. Indeed, if this were not the case, by putting $K \equiv X$ and $Y \equiv R$ in (22), we would have for each $K \in \mathbf{U}$, $R \subset G(K) \subset K$, whence $R \subset \bigcap \mathbf{U}$ and $R = \bigcap \mathbf{U}$, contrary to the hypothesis.

Now, $K \subset R$ and $R \not\subset \bigcap \mathbf{U}$ imply, using $R \equiv X$ and $Y \equiv \bigcap \mathbf{U}$ in (22), that $\bigcap \mathbf{U} \subset G(R)$, whence $G(R) \in \mathbf{R}$. Therefore $R \in \mathbf{R}$ implies $G(R) \in \mathbf{R}$.

[**Inner limit case:**] Finally: $\emptyset \neq \mathbf{X} \subset \mathbf{R}$ implies $\bigcap \mathbf{X} \in \mathbf{R}$. Indeed, if all of the elements of $\mathbf{X}$ contain $\bigcap \mathbf{U}$, so does their intersection $\bigcap \mathbf{X}$. Otherwise, there would be some member of $\mathbf{X}$ that is a subset of $G(\bigcap \mathbf{U})$, so, *â fortiori*, $\bigcap \mathbf{X} \subset G(\bigcap \mathbf{U})$. Thus in all cases $\bigcap \mathbf{X} \in \mathbf{R}$.

[**Inner conclusion:**] According to Theorem I,

$$\mathbf{R} \;=\; \mathbf{M}(A). \tag{32}$$

The result of this is that $\bigcap \mathbf{U} \in \mathbf{K}$. Indeed, let $X \in \mathbf{M}(A)$ such that $\bigcap \mathbf{U} \subset X$. If there is also an element of $K \in \mathbf{U}$ such that $K \subset X$, one can of course put $\bigcap \mathbf{U}$ for K in (22). Otherwise all of the elements $K \in \mathbf{U}$ contain $\mathbf{X}$ by (22), so $\bigcap \mathbf{U} \supset X$, whence $\bigcap \mathbf{U} = X$ and, by (32), substituting $\bigcap \mathbf{U}$ for K in (22), we obtain the formula $\bigcap \mathbf{U} \in \mathbf{K}$.

[**Outer conclusion:**] Having established the conditions 8–10 for Theorem I, the result is that any element of $\mathbf{M}(A)$ is a K that satisfies (22). However, putting $X \equiv K$ in (22), we immediately deduce that $\mathbf{M}(A)$ is a decreasing family of subsets. Moreover, if X is any element of $\mathbf{M}(A)$, there is no element of $\mathbf{M}(A)$ that is different from both X and $G(X)$ while being a subset of X and containing $G(X)$. (33)

In other words:

$$G(X) \quad \text{is either equal to} \quad X \quad \text{or is its immediate successor.} \tag{34}$$

Corollary I. $\mathbf{M}(A)$ is a well ordered decreasing family of sets.

In other words, any nonempty subfamily of $\mathbf{M}(A)$ has a largest member. This can be written as follows:

$$\emptyset \neq \mathbf{X} \subset \mathbf{M}(A) \quad \text{implies} \quad \bigcup \mathbf{X} \in \mathbf{X}. \tag{35}$$

Note that the phrase "well ordered descending family of sets" makes no appeal to the general notion of order: this phrase is defined by the condition (35) [Kur21].

Proof. Indeed, if $A \in \mathbf{X} \subset \mathbf{M}(A)$ then by (23) A itself is the largest element of $\mathbf{X}$. Otherwise, let $\mathbf{P}$ be the family of all of the members of $\mathbf{M}(A)$ that contain those of $\mathbf{X}$. If $\bigcap \mathbf{P} \in \mathbf{X}$ then $\bigcap \mathbf{P}$ is the largest element of $\mathbf{X}$; otherwise $\bigcap \mathbf{P} \notin \mathbf{X}$ and so $G(\bigcap \mathbf{P}) \in \mathbf{X}$ and, by (34), $G(\bigcap \mathbf{P})$ contains all the other elements of $\mathbf{X}$. Hence, in either case, $\bigcup \mathbf{X} \in \mathbf{X}$. QED

Symmetrically, we have

Corollary I′. $\mathbf{N}(A)$ is a well-ordered increasing family of sets and

$$\emptyset \neq \mathbf{X} \subset \mathbf{N}(A) \quad \text{implies} \quad \bigcap \mathbf{X} \in \mathbf{X}. \tag{36}$$

Now consider family $\mathbf{D}$ of all of the differences $M \setminus G(M)$ for $M \in \mathbf{M}(A)$. We claim that

$$A \;=\; P(A) \;+\; \bigcup \mathbf{D}, \tag{37}$$

where $P(A)$ is, as usual, the smallest member of the family $\mathbf{M}(A)$.

To prove this formula it is enough to show that, for any element $p \in A \setminus P(A)$, there some element $M(p) \in \mathbf{M}(A)$ such that

$$p \in M(p) \quad \text{and} \quad p \notin G\big(M(p)\big). \tag{38}$$

Indeed, among those M that do not contain p, there is, by Corollary I, the largest, so write M_0 for it. By hypothesis, $p \in A$, so $M_0 \neq A$. Also, the set M_0 cannot be the intersection of larger sets than itself that belong to $\mathbf{M}(A)$, for, if they contained p, so would M_0 itself. By definition of $\mathbf{M}(A)$, there is therefore some $M(p)$ such that $M_0 = G\big(M(p)\big)$, which proves that the formula (38) is valid.

In addition, for any given p there is exactly one set $M(p)$ that satisfies condition (38). Indeed, by Theorem III, the sets M may be divided into two families: the super-sets of $M(p)$ and the subsets of $G\big(M(p)\big)$. The set $M(p)$ is the smallest of the first family and $G\big(M(p)\big)$ is the largest in the second.

In the particular case where the family $\mathbf{M}(A)$ consists of a sequence $M_1, M_2, \ldots, M_n, \ldots$, the formula (37) gives decomposition of A into two separate parts

$$A \;=\; P(A) \,\cup\, \bigcup_{n=1}^{\infty} \big(M_n \setminus G(M_n)\big). \tag{39}$$

FN1) The position that the elements of $\mathbf{M}(A)$ occupy in the sequence $\{M_n\}$ obviously does not depend on their order of decreasing in $\mathbf{M}(A)$. (????)

In addition, for each element $p \in A \setminus P(A)$ there is only one number $n(p)$ such that

$$p \in M_{n(p)} \quad \text{and} \quad p \notin G(M_{n(p)}). \tag{40}$$

Similarly, there are symmetrical formulae to (37–40) for $\mathbf{N}(A)$ instead of $\mathbf{M}(A)$.

Note: This is actually only a third of the paper: the remainder considers applications. I will translate it at a later date.

Translator's note

Kuratowski credits his Theorem III and its Corollaries to Gerhard Hessenberg [Hes08, p.127]. (I tried to translate this paper with the help of a German mathematician, looking for this proof, but we found the notation impenetrable.) Hessenberg collaborated with Ernst Zermelo, who used this Theorem as the engine of his second proof of the Well-Ordering Theorem [Zer08a]: see the last paragraph of [vH67, p.184]. It became known (of course many decades later) as the Bourbaki–Witt Theorem [Bou49, Wit51] and Walter Felscher [Fel62] compared the many published versions of it.

There are several strategies for the proof. Kuratowski's makes heavy use of Excluded Middle, but Todd Wilson has given a constructive version [Wil01]. Even so, the resulting notion of well-ordering just says that every inhabited subset has a least element, as in Cantor's condition, which is not enough for intuitionistic induction. Nevertheless, it means that the Well-Ordering Theorem invokes Choice at the beginning and Excluded Middle at the end, but the key argument uses neither of them.

The principal message of Kuratowski's paper is that it is not *necessary* to use ordinals to prove theorems of the pattern that he describes. But we may go further than this to say that the naïve use of them is not even *sufficient* (valid). This is the reason why we have introduced the distinction between "families" and "classes" in the translation.

Because of the Burali-Forti paradox [BF97], the ordinals do not form a set definable in Zermelo's axioms. The class $\mathbf{A}(A)$ is a sub-*collection* of $\mathcal{P}(A)$, but to define it as a *set* (*family* in our usage) using the Axiom of Separation requires *unbounded existential quantification* over the class of ordinals, or equivalently an Axiom of Collection, *i.e.* that the image of a class under a function to a set is again a set. (One may argue instead that Zermelo's notion of a "definite property" in his Axiom of Separation is so vague that it allows $\mathbf{A}(A)$.)

In order to make $\mathbf{A}(A)$ into a legitimate set (a family $\mathbf{Z}$) without such trickery, we must say *when to stop* iterating over the ordinals. One way of doing that is Friedrich Hartogs' construction [Har15], but Theorem III provides a neater way. That is to say that it shows explicitly that the family $\mathbf{M}(A)$ is a form of the class $\mathbf{A}(A)$ restricted to a particular ordinal.

Another issue with the purported proof using ordinals is that the recursion that is needed to satisfy formulae (2–5) depends on a Theorem that was only proved later in the decade in which Kuratowski had written the present paper, namely by John von Neumann [vN28, Section 3].

References

[BF97] Cesare Burali-Forti. Una questione sui numeri transfiniti. *Rendiconti del Circolo matematico di Palermo*, 11:154–164, 1897. English translation in [vH67], pages 104–111.

[Bou49] Nicolas Bourbaki. Sur le théorème de Zorn. *Archiv der Mathematik*, 2:434–7, 1949. English translation at `www.paultaylor.eu/trans/`.

[Can15] Georg Cantor. *Contributions to the Founding of the Theory of Transfinite Numbers*. Open Court, 1915. Translated and with a historical introduction by Philip Jourdain; republished by Dover, 1955; `archive.org/details/contributionstot003626mbp`.

[Can32] Georg Cantor. *Gesammelte Abhandlung mathematischen und philosophischen Inhalts*. Springer-Verlag, 1932. Edited by Ernst Zermelo.

[Ded88] J. W. Richard Dedekind. *Was Sind und was Sollen die Zahlen?* Braunschweig, 1888. English translation in [Ded94].

[Ded94] J. W. Richard Dedekind. The nature and meaning of numbers. In *Essays on the Theory of Numbers*. 1894. Dover (1960).

[Fel62] Walter Felscher. Doppelte Hülleninduktion und ein Satz von Hessenberg und Bourbaki. *Archiv der Mathematik*, 13:160–5, 1962. English translation at `www.paultaylor.eu/trans/`.

[Har15] Friedrich Hartogs. Über das Problem der Wohlordnung. *Mathematische Annalen*, 76:438–443, December 1915. English translation at `www.paultaylor.eu/trans/`.

[Hau14] Felix Hausdorff. *Grundzüge der Mengenlehre*. 1914.

[Hes08] Gerhard Hessenberg. Kettentheorie und Wohlordnung. *Journal für die reine und angewandte Mathematik (Crelle's Journal)*, 135:81–133, 1908.

[KK21] Bronisław Knaster and Kazimierz Kuratowski. Sur les ensembles connexes. *Fundamenta Mathematicae*, 2(1):206–255, 1921.

[KS21] Kazimierz Kuratowski and Wacław Sierpiński. Le théorème de Borel–Lebesgue dans la théorie des ensembles abstraits. *Fundamenta Mathematicae*, 2(1):172–178, 1921.

[Kur21] Kazimierz Kuratowski. Sur la notion de l'ordre dans la théorie des ensembles. *Fundamenta Mathematicae*, 2(1):161–171, 1921.

[Kur22] Kazimierz Kuratowski. Une méthode d'élimination des nombres transfinis des raisonnements mathématiques. *Fundamenta Mathematicae*, 3:76–108, 1922.

[Leb05] Henri Léon Lebesgue. Sur les fonctions representables analytiquement. *Journal de Mathématiques Pures et Appliquées*, 1:139–216, 1905.

[Lin05] Ernst Lindelöf. Remarques sur un théorème fondamental de la théorie des ensembles. *Acta Mathematica*, 29(none):183 – 190, 1905.

[Sie20] Wacław Sierpiński. Une démonstration du théorème sur la structure des ensembles de points. *Fundamenta Mathematicae*, 1(1):1–6, 1920.

[vH67] Jan van Heijenoort. *From Frege to Gödel: a Source Book in Mathematical Logic, 1879–1931*. Harvard University Press, 1967.

[vN28] John von Neumann. Über die Definition durch transfinite Induktion und verwandte Fragen der allgemeinen Mengenlehre. *Mathematisches Annalen*, 99:373–393, 1928.

[Wil01] Todd Wilson. An intuitionistic version of Zermelo's proof that every choice set can be well-ordered. *Journal of Symbolic Logic*, 66:1121–6, 2001.

[Wit51] Ernst Witt. Beweisstudien zum Satz von M. Zorn. *Mathematische Nachrichten*, 4:434–8, 1951.

[YY06] William Henry Young and Grace Chisholm Young. *The theory of sets of points*. Cambridge University Press, 1906.

[Zer04] Ernst Zermelo. Beweis, daß jede Menge wohlgeordnet werden kann. *Mathematisches Annalen*, 59:514–6, 1904. English translation in [vH67].

[Zer08a] Ernst Zermelo. Neuer Beweis für die Möglichkeit einer Wohlordnung. *Mathematisches Annalen*, 65:107–128, 1908. English translation in [vH67], pages 183–198.

[Zer08b] Ernst Zermelo. Untersuchungen über die Grundlagen der Mengenlehre I. *Mathematische Annalen*, 59:261–281, 1908. English translation in [vH67], pages 199–215.