A Method for Eliminating Transfinite Numbers from Mathematical Arguments

by Casimir (Kazimierz) Kuratowski, translated by Paul Taylor

1922 and 2025

This translation is dedicated to the memory of my friend Marek Zawadowski (1960–2024), a great-grandstudent of Kuratowski with equally deep roots in the city of Warsaw.

It is not yet complete, but I will translate more of the "applications" in due course: It is difficult to track down all of the references and tedious to type the LaTeX.

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 [Georg] 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 [Henri] Lebesgue [Leb05], [Ernst] Lindelöf for the Cantor–Bendixson Theorem [Lin05], [Wacław] Sierpiński [Sie20], the [William and Grace] Youngs for the class of derivatives of any order of a given set of points [YY06, appx] and [Ernst] Zermelo for the Well-Ordering Theorem [Zer04, Zer08a], but using different methods in each particular case. Lebesgue shows without ordinals an important theorem on Class 1 functions that [René-Louis] Baire had proved in his thesis [Bai99] using these numbers.

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 [Richard] 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 [Gerhard] 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,

[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".

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.

[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 just 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

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

[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.

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

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

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

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 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.

[Replacing the classes]

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

[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 $\setminus$.

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

- [Felix] Bernstein's results about sets without perfect parts [Ber08],
- in [Paul] Mahlo's about the number of "homoïes" [Mah11],
- in the article by [Wacław] Sierpiński and myself on classes (**L**) [KS21] and
- in the one by [Bronisław] Knaster and myself on connected sets [KK21]

is similar to their role in the reasoning on family **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$'$.

[Induction]

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 $\mathbf{A}(A)$ is 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 $X \subset A$.]

[Fixed Points]

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).

[Well-ordering]

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

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

[Difference sets]

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

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

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 $\mathbf{M}(A)$.

1 Applications: Zermelo's theorem

Zermelo himself proved his famous theorem [that every set can be well ordered, assuming his Axiom of Choice] without using ordinals [Zer08a]. To do this using the method above, it is enough to put, for any given set E,

$$A \;\equiv\; E \quad \text{and} \quad G(X) \;\equiv\; X \setminus F(X),$$

where $F(X)$ denotes a function that, to any non-empty subset $X \subset E$, assigns a singleton $F(X) \equiv \{x\} \subset X$. That such a function exists is the Axiom of Choice.

For any two distinct elements, we write $p \prec q$ if $M(p) \supset M(q)$, where $M(p)$ and $M(q)$ satisfy (38). This is a well-ordering on E. Indeed, we have $P(A) = \emptyset$ by (12). We can therefore match each $p \in E$ with a particular $M(p)$. Also, when $p \neq q$, $M(p) \neq M(q)$, because

$$M(p) \setminus G(M(p)) \;=\; F(M(p)) = \{p\}.$$

According to Corollary I, the set E is well ordered.

2 Saturated and irreducible sets; Brouwer's theorem

[The section is about what became known as *Zorn's Lemma* [Zor35]. To be fair, Max Zorn denied credit for it when interviewed by Paul Campbell for his history of the topic [Cam78], which traced the first occurrence to Felix Hausdorff [Hau06] and also the authors mentioned below.]

[Zygmunt] Janiszewski [Jan12] defined a set to be

- *saturated* with respect to a given property if it satisfies it but it is not a proper subset of any other set that does so [*i.e.* maximal], and

- *irreducible* if it has it but no proper subset of it does so [*i.e.* minimal].

These concepts have been considered several times, especially in Analysis Sitūs [point–set topology], but the following theorem concerns sets whose elements are completely arbitrary in nature.

> (41) Let E be a set satisfying $\mathbf{R}(E)$ such that, for any descending well ordered family of subsets having the property, their intersection does so too. Then there is some subset that is irreducible with respect to $\mathbf{R}$.

We will prove this theorem using the properties of the family $\mathbf{M}(A)$ defined on page 3. Let $A \equiv E$ and let $G(X)$ be some function of $X \subset E$ such that

- $G(X) \equiv Y$ for some proper subset $Y \subset X$ with $\mathbf{R}(Y)$, but
- $G(X) \equiv X$ if there is no such Y.

That such a function exists follows immediately from the axiom of choice.

Then we claim that $P(A)$ is the required irreducible set.

Let $\mathbf{P} \subset \mathbf{M}(A)$ consist of the elements that have the property $\mathbf{R}$. It is a family $\mathbf{Z}$ satisfying conditions (7–10). Indeed, if $X \in \mathbf{P}$ then $G(X) \in \mathbf{P}$ by the definition of $G(X)$. Moreover, if $\mathbf{X} \subset \mathbf{P}$ then by Corollary I X is a descending well ordered sequence of subsets, so by hypothesis, $\bigcap \mathbf{X} \in \mathbf{P}$.

According to the theorem I, any element of $\mathbf{M}(A)$ has the property $\mathbf{R}$. In particular, $P(A)$ has it.

Further, by (12), $P(A) = G(P(A))$, which means precisely that no proper subset of $P(A)$ satisfies $\mathbf{R}$. Therefore $P(A)$ is irreducible with respect to $\mathbf{R}$.

In a similar way it is shown that

> (42) Let E be a set satisfying $\mathbf{R}(E)$ such that, for any increasing well ordered sequence of subsets having the property, their union does so too. Then there is some subset that is irreducible with respect to $\mathbf{R}$.

An immediate consequence of this is a theorem of [L.E.J.] Brouwer [Bro10]:

> Let E be a closed set of points that satisfies $\mathbf{R}$ and such that, for any simply infinite decreasing sequence of subsets each satisfying $\mathbf{R}$, the intersection does so too. Then one can reduce the set E by a countably infinite number of operations to a closed set that also satisfies $\mathbf{R}$, but for which no proper closed subset does so.

Indeed, it is sufficient to replace $\mathbf{R}$ in (41) by the property of being a *closed* subset satisfying $\mathbf{R}$. Now, $\mathbf{M}(A)$ is in this case a decreasing well ordered sequence of closed sets; it is therefore finite or countable. We therefore need at most $\aleph_0$ operations to obtain its last element, $P(A)$. This is indeed the required subset, since it is irreducible with respect to the property of being a closed set satisfying $\mathbf{R}$.

[Ludovic] Zoretti introduced in Analysis Sitūs the important notion of an *irreducible continuum* between two points [Zor10], that is, a continuum that is irreducible with respect to the property of being a continuum with two given points a and b as elements.

As Janiszewski showed, every bounded continuum containing points a and b contains an irreducible continuum between these points. This result can be deduced directly from (41), since the intersection of any infinite decreasing sequence continua is a continuum or just a point [Jan12, p104].

Janiszewski's proof [Jan10] may be seen as a special case of Schema I. [Stefan] Mazurkiewicz proved the same result by a quite different method, without ordinals [Maz10]. He gave another proof in [Maz19].

3 Fréchet's abstract families

Let E be any set and suppose that a function $f(e_1, e_2, ...e_n, ...)$ is defined on E such that

- if $e_1 = e_2 = \cdots = e_n = \cdots = e$ then $f(e_1, e_2, ...e_n, ...) = e$; and
- if $f(e_1, e_2, ..., e_n, ...) = g$ then for $n_1 < n_2 < n_3 < \cdots$ we have $f(e_{n_1}, e_{n_2}, \ldots) = g$;
- although there may be sequences $\{e_n\}$ for which $f(e_1, e_2, ...e_n, ...)$ is not defined.

We say, following Fréchet [Fré06], that E is a *class* (**L**) and that $f(e_{n_1}, e_{n_2}, \ldots)$ is the *limit* of the sequence $\{e_n\}$.

For any subset $X \subset E$ we denote by $G(X)$ the set of all limits of sequences in X, so $X \subset G(X)$.

The sequence $\{A_\alpha\}$ is defined naturally as follows:

Let $A \subset E$ be any subset. Put

- $A_0 = A$;
- if A_α is defined then $A_{\alpha+1} = G(A_\alpha)$;
- if α is a limit ordinal and A_ξ is defined for all $\xi < \alpha$ then $A_\alpha = \bigcup_{\xi<\alpha} A_\xi$.

We immediately see that this definition is an application of Schema I'. We may therefore replace it with the definition of the family $N(A)$, as established earlier. The function $G(X)$ obviously satisfies condition (20); therefore according to Theorem II', $S(A)$ is the smallest superset of A that satisfies equation (15). In other words, $S(A)$ is the smallest *closed* superset of A. We can therefore consider $S(A)$ as a generalisation of the set $\overline{A}$ in Analysis Situs [*i.e.* the closure in point–set topology].

We will now show that the family of all elements of $N(A)$ that precede a given element N is at most countable, unless N is identical to $S(A)$, which is the last member of the family $N(A)$ ordered by size (Corollary I').

In fact, suppose that N_0 is the smallest element of $N(A)$ preceded by uncountably many elements of this family and also that $N_0 \neq S(A)$. However, since N_0 clearly has no immediate predecessor, it is the union of all of its predecessors.

Since $N_0 \neq S(A)$, there is some e with $e \in G(N_0) \setminus N_0$, so let $\{e_n\}$ be a sequence in N_0 that "tends" to e. We denote by N_n for $n \geq 1$ the first set belonging to the family $\mathsf{N}(A)$ that has e_n as an element.

Hence $N_n \subset N_0$ and $N_n \neq N_0$. Therefore there are at most $\aleph_0$ elements of $N(A)$ that precede N_n for some $n \geq 1$. The same applies to the union $I = \bigcup_{n=1}^{\infty} N_n$ and therefore $I \neq N_0$. But since all of the e_n belong to I, their limit e belongs to $G(I)$.

This is in contradiction with the formula $e \in G(N_0) \setminus N_0$, since $G(I) \subset N_0$.

This proves our theorem.

In terms of the theory of ordinals, we say that the order type of the family $N(A)$ is at most $\Omega + 1$.

4 Dérivés and coherences

I don't know what "dérivés" are in modern mathematical terminology.

As we said on p. 78, the definition of the class of coherences of any order of a set of points A is obtained from Schema I by setting

$$G(X) \equiv X \cap X' \tag{43}$$

for any subspace X of a Euclidean space E.

The condition that E be a *Euclidean* space is not essential for the reasoning in this section. It remains valid when E denotes, for example, a normal class $(\mathcal{E})$. Thus, the proof of the Cantor–Bendixson theorem extended to these classes by Fréchet can be expressed without ordinals [Fré10].

The function $G(X)$ being subject to condition (43), we will call $\mathsf{M}(A)$ the class of *coherences* of any order of A. The family $\mathsf{M}(A')$ will be called the family of *derivatives* of any order of A.

This definition differs little from that of Young, which we have mentioned on p. 77. However, it should be noted that Young demonstrates the property of the class of derivatives of being a class of decreasing sets by relying on several specific properties of derivatives (Theorems 2, 4, 5) that are not needed when applying Theorem III. The only property of derivatives that needs to be considered is that $G(X) \subset X$, which follows immediately from formula (43).

To convince oneself that the family $\mathsf{M}(A')$ coincides with the class of derivatives of order $\alpha \geq 1$ of A, one need only note that $G(X) = X'$ when X is closed.

The fundamental properties of derivatives and coherences can be easily deduced from the general theorems on the familyes M that have been established previously. We will consider the best-known properties.

It follows from Theorem I that any element M of the family $\mathsf{M}(A)$ is closed in A. (X is said to be *closed in* Y when X is the intersection of Y with a closed subset.) Indeed, if X is closed in A, $G(X)$ is also closed, since $G(X)$ is closed in X by (43); moreover, the intersection of closed sets in A is also closed in A.

In particular, when A is closed, the elements of $\mathsf{M}(A)$ are also closed. However, since the derivative of any set is closed, we can deduce that the derivatives of any order, *i.e.* all elements of $\mathsf{M}(A')$, are closed, regardless of the set A.

According to Corollary I, $\mathsf{M}(A)$ is a well-ordered family of decreasing sets. Since $G(X)$ is always closed in X, we can find in each element $M \neq P(A)$ of $\mathsf{M}(A)$ a point that is not a limit point of the elements of $\mathsf{M}(A)$ following M. We can therefore arrange the elements of $\mathsf{M}(A)$ in a sequence: $M_1, M_2, \ldots, M_n, \ldots$. This can be done in a specific way according to the rule given by Sierpiński.[4]

According to (39), we have the decomposition:

$$A \;=\; P(A) \;+\; \bigcup_{n=1}^{\infty} \big(M_n \setminus G(M_n)\big), \tag{44}$$

where $P(A)$ denotes the *last* coherence. According to (43), the differences $M_n \setminus G(M_n)$ are isolated sets.

Let us now note that, when $X \subset Y$, we have $X' \subset Y'$, whence $G(X) \subset G(Y)$. According to theorem II, the set $P(A)$ is therefore the largest subset of A that satisfies the equation $Z = G(Z)$. By (43), this means precisely that $P(A)$ is the largest subset of A that is dense-in-itself. It follows that the union $\bigcup_{n=1}^{\infty} \big(M_n \setminus G(M_n)\big)$ does not contain any dense-in-itself set, and is therefore a sparse set.

When A is closed, $P(A)$ is also closed because $P(A)$ is always closed in A. Since, on the other hand, $P(A)$ is always dense-in-itself, we can deduce that if A is closed then $P(A)$ is perfect.

Thus, for the case of closed set A, formula (44) provides us with a decomposition of A into two disjoint sets, the first of which is perfect and the second sparse.

[4]The citation is just Bull. Acad. Polonaise des Sciences July 1921.

5 Sparse sets and the Cantor—Bendixson theorem

According to the theorem of Cantor and Bendixson, every closed set consists of a perfect set and a finite or countable set. This theorem was first established using transfinite numbers.

Later, Lindelöf proved it using the notion of a *condensation point*, without resorting to transfinite numbers. However, it relied on the axiom of choice because he assumed that the union of countably many countable sets is countable [Sie18]. This is also used in [You03]. So it was desirable to provide a proof that did not involve either the transfinite or the axiom of choice, as Sierpiński recently did [Sie20].

This can be done by applying our general method to the classical argument in *Acta Mathematica*, volume 2. As we have just seen, Equation (44) expresses the closed set A as the union of a perfect set and a sequence of isolated sets. To deduce the Cantor–Bendixson theorem from this, we need only show how to arrange the points of an isolated set in a finite finite or infinite sequence.

Let I be any isolated set and let $R_1, R_2, \dots R_i, \dots$ be the sequence of spheres with rational centres and coordinates. For each point $p \in I$, there is a number $i(p)$ such that p is the only point in the set $I \cap R_{i(p)}$; suppose that $i(p)$ is the smallest number with this property. We have thus assigned an index $i(p)$ to every point $p \in I$; moreover, the same index clearly cannot be assigned to two distinct points of I. The set I is therefore arranged in a sequence, as required.

Sierpiński's proof can be expressed using the method that we have described. As we said at the end of Section 4, any closed set can be decomposed into a perfect set and a sparse set. Moreover, as Sierpiński shows, we can decompose closed sets into perfect and sparse sets, independently of the axiom of choice and without involving classes of consistency and derivatives. The problem is therefore reduced to showing that any sparse set is finite or countable.

Let E be a sparse set. If $\emptyset \neq X \subset E$ is sparse, it follows that there is a number $i(X)$ such that X has only one point in common with the sphere $R_{i(X)}$ of the sequence $\{R_i\}$ considered earlier. Suppose that $i(X)$ is the smallest number with this property.

Put $F(X) = X \cap R_{i(X)}$. The set $F(X)$ consists of a single point, unless X is empty, in which case, we put $F(\emptyset) = \emptyset$.

Put

$$G(X) \equiv X + F(E \setminus X) \tag{45}$$

$$A \equiv F(E) \tag{46}$$

and consider the family $N(A)$.

Since the union $S(A)$ of all elements of $N(A)$ satisfies equation (15), we have $S(A) = E$ by (45). Therefore for every point $p \in E \setminus A$ there is a set N_p such that such that $p \in G(N_p) \setminus N_p$. By (45), $\{p\} = F(E \setminus N_p)$. Let us therefore set

- $n(a) = \emptyset$ when $a \in A$ and
- for $p \in (E \setminus A)$ let $n(P) = i(E \setminus N_p)$.

We claim that the inequality $p \neq r$ implies $n(p) \neq n(r)$. Indeed, we have $N_p \neq N_r$ and if $n(p) = n(r)$ we have $F(E \setminus N_p) = (E \setminus N_p) \cap R_{n(p)}$ and $F(E \setminus N_r) = (E \setminus N_r) \cap R_{n(p)}$.

But since the family $N(A)$ is a family of increasing sets, we can put $N_p \subset N_r$, which gives $F(E \setminus N_p) \supset F(E \setminus N_r)$. But the last inclusion is absurd, since $F(E \setminus N_p) = \{p\}$ and $F(E \setminus N_r) = \{r\}$, where p and r are distinct.

It is thus established that $n(p) \neq n(r)$ and so the elements of E are enumerated.

6 Residues and sets that are both F_σ and G_δ

The concept of residue, like that of consistency, leads to an important decomposition of sets of points. Following Hausdorff [Hau14, p. 281], for any set X, we call the subset

$$G(X) \equiv X \cap \overline{\overline{X} \setminus X}$$

the *residue* of X, where $\overline{X}$ denotes the subset consisting of X and its limit points.

For a given set A, put

- $A_0 \equiv A$;
- $S_{\alpha+1} \equiv G(A_\alpha)$; and
- $A_\alpha \equiv \bigcap_{\xi<\alpha} A_\xi$ when α is a limit ordinal.

Then we may consider the residues of A of order α.

To do this without using ordinals, we need only consider the class $\mathbf{M}(A)$, where E denotes Euclidean space.

By (47), $G(X)$ is closed in X. It is easy to see that all of the elements of $\mathbf{M}(A)$ are closed subsets of A and that the class $\mathbf{M}(A)$ is formed by a sequence $M_1, M_2, \ldots$. By (39), we have

$$A = P(A) \cup \bigcup_{n=1}^{\infty} (M_n \setminus F(M_n)),$$

where $P(A)$ denotes the last residue of A.

Many interesting properties of residues were established by Hausdorff in his book and it is not difficult to eliminate ordinals from his proofs. In particular, we show that

1. if $P(A) = \emptyset$ then $P(E \setminus A) = \emptyset$;
2. the last residue of the set $\bigcup_{n=1}^{\infty} \big(M_n \setminus G(M_n)\big)$ is empty;
3. if A is both F_σ and G_δ then $P(A) = \emptyset$; and
4. among the closed subsets of A, $P(A)$ is the largest that satisfies the equality $X = G(X)$.

In this, any subset that is the union of a sequence of closed sets is called F_σ. The complements of such subsets, *i.e.* the intersections of sequences of open sets, are called G_δ.

Now note that

$$X \setminus G(X) = X \setminus \overline{\overline{X} \setminus X} = \overline{X} \setminus \overline{\overline{X} \setminus X},$$

which shows that the set $X \setminus G(X)$ is the difference of two closed sets. In other words, it is the intersection of a closed subset with an open one. The subset $X \setminus G(X)$ is therefore the union of a (well-defined) sequence of closed sets.

Consequently, if $P(A) = \emptyset$ then we deduce from (48) that A is an F_σ. Moreover, A can be expressed in a canonical way as the union of a sequence of closed sets. The equality $P(A) = \emptyset$ therefore implies that A is an *effective* F_σ. However, as Sierpiński remarked, we do not know how to extend this to the case of a set that is just F_σ or just G_δ [Sie21, p. 114].

It follows, according to (1) and (3), that: *for a set to be both F_σ and G_δ, it is necessary and sufficient that its last residue be empty.*

Furthermore: *any set that is both F_σ and G_δ effectively so.*

Taking into account (2) and (4) we deduce: *any subset A can be decomposed in a unique way into two disjoint subsets, one of which is closed in A and equal to its residue, and the other of which is both F_σ and G_δ.*

Formula (48) in particular represents this decomposition.

We conclude this section with a theorem that Young proved using ordinals [YY06, p. 65].

It is easy to see that if p is an isolated point of a set X, the residue $G(X)$ does not contain this point. Furthermore, no sparse (non-empty) set satisfies the formula $X = G(X)$. Similarly, no subset of a sparse set A satisfies this formula either: however, according to (12), $P(A) = G\big(P(A)\big)$, so $P(A) = \emptyset$. As we indicated earlier, this last equality implies that A is both F_σ and G_δ.

This brings us to the theorem: *Every sparse set is* G_δ.

7 Borel classes

The classication of *measurable* sets in the sens of Borel leads to the consideration of two transfinite sequences of classes:

$$F,\ F_\sigma,\ F_{\sigma\delta},\ F_{\sigma\delta\sigma},\ \ldots F_{(\alpha)} \tag{49}$$

$$G,\ G_\delta,\ G_{\delta\sigma},\ G_{\delta\sigma\delta},\ \ldots G_{(\alpha)} \tag{50}$$

for ordinals $alpha < \Omega$ [Hau14, p. 304].

In this section we define these sequences and indicate their fundamental properties without recourse to the transfinite. In fact, we will only consider the first sequence, since the second is symmetric with it,

Let E the class of all of the sets of points (in a Euclidean space) and A the class of closed subsets.

For any subclass $X \subset E$, let $R(X)$ be the class composed of the members of X and the intersections of countably many subsets that are membes of X; in other words, if Y is any finite or countable subclass of X, $\bigcap Y \in R(X)$.

In just the same way, $T(X)$ with denote the class of members of X and countable unions of them. We put

$$\begin{cases} \text{when} \quad X = R(X), & G(X) = T(X), \\ \text{when} \quad X \neq R(X), & G(X) = R(X). \end{cases} \tag{51}$$

The function $G(X)$ clearly satisfies the inclusion $X \subset G(X)$. $\mathsf{N}(A)$ is of course the family of all of the Borel classes F.

It is easy to see that $R(X) \cup T(X) \subset GG(X)$. On the other hand, the inclusion $X \subset Y$ implies $R(X) \subset R(Y)$ and $T(X) \subset T(Y)$, so

$$G(X)\ \subset\ R(Y) \cup T(Y)\ \subset\ GG(Y).$$

But $GG(Y) \subset \bigcup \mathsf{N}(Y)$; thus $X \subset Y$ implies $G(X) \subset \bigcup \mathsf{N}(Y)$ and from (21) the class $S(A)$ of all measurable subsets (B) is the smallest class Z that satisfies the conditions

$$A\ \subset\ Z \qquad \text{and} \qquad Z\ =\ G(Z).$$

In other words, (52) $S(A)$ is the smallest class that has as elements all of the closed subsets and is closed under countable unions and intersections, *cf.* [Hau14, p. 305].

We will say that a Borel class N is *even* if $N = R(N)$ and otherwise *odd*. The evident identities $RR(X) = R(X)$ and $TT(X) = T(X)$ show that the even and odd classes alternate.

Now let p be a measurable set (B) that isn't closed and let $N(p)$ be the largest Borel class to which p doesn't belong. Then $p \in G(N(p))$ and $p \notin N(p)$. Therefore there exists a sequence

$$p_1,\ p_2,\ p_3,\ \ldots \tag{53}$$

of sets that are members of $\mathsf{N}(p)$ such that

$$\text{either} \qquad p \ = \bigcup_{n=1}^{\infty} \qquad \text{or} \qquad p \ = \bigcap_{n=1}^{\infty} \tag{54}$$

according as the class $\mathsf{N})p)$ is even or odd.

We will show by induction that

$$\text{if } N' \text{ is a Borel class, the conditions } p \in N \text{ and } r \in N \text{ entail } (p \cup r) \in N \text{ and } p \cap r) \in N. \tag{55}$$

Let P be the family of Borel classes that satisfy condition (55).

So, since the umion and the intersection of two closed subsets are closed, we have

$$A \in \mathsf{P} \tag{56}$$

Also, suppose that $X \in \mathsf{P}$. We claim that $G(X) \in \mathsf{P}$. Indeed, if p and r are any two members of $G(X)$, there are two sequences of members of X, $\{p_n\}$ and $\{r_n\}$ such that either

$$p \ = \bigcup_{n=1}^{\infty} p_n \qquad \text{and} \qquad r \ = \bigcup_{n=1}^{\infty} r_n \tag{57}$$

or

$$p \ = \bigcap_{n=1}^{\infty} p_n \qquad \text{and} \qquad r \ = \bigcap_{n=1}^{\infty} r_n \tag{57}$$

according as the class X is even or odd.

In the first case we have $p \cup r = \bigcup_{n=1}^{\infty}(p_n \cup r_n)$ and since, by hypothesis, $(p_n \cup r_n) \in \mathsf{X}$, we deduce that $(p \cup r) \in G(X)$.

On the other hand,

$$p \cap r \ = \ \Big(\bigcup_{n=1}^{\infty} p_n\Big) \ \cap \ \Big(\bigcup_{m=1}^{\infty} r_m\Big) \ = \ \bigcup_{n=1, m=1}^{\infty} p_n \cap r_m,$$

which implies, since $(p_n \cap r_m) \in \mathsf{X}$, that $(p \cap r) \in G(\mathsf{X})$.

Thus, condition (55) implies that $G(X) \in \mathsf{P}$. In an entirely analgous way, we can show the same for condition (58). So we can assert that

$$X \in \mathsf{P} \qquad \text{entails} \qquad G(X) \in \mathsf{P}. \tag{59}$$

We will show that

$$X \subset \mathsf{P} \qquad \text{entails} \qquad \bigcup X \in \mathsf{P}. \tag{60}$$

Let p and r be any two members of $\bigcup X$. Since the class X is an increasing sequence of subset, there is an element $V \in X$ such that $p, r \in V$. By definition of X, $(p \cup r), (p \cap r) \in V$ and therefore $(p \cup r), (p \cap r) \in \bigcup X$, which justifies formula (60).

From formulae (56), (59) and (60) we deduce (55) from Theorem I'.

We showed previously that every set p that isn't closed and measurable in the sense of Borel is the union or the intersection of a sequence $\{p_n\}$ of members of $N(p)$. Applying this to (55), we may suppose that this sequence obeys one of the two conditions

$$p_1 \subset p_2 \subset p_3 \subset \cdots \qquad \text{or} \qquad p_1 \supset p_2 \supset p_3 \supset \cdots,$$

depending on whether p is the union or the intersection of the $\{p_n\}$.

By an analogous method, we could eliminate ordinals from the proofs of other theorems about Borel classes. Thus, for example, there is no difficulty in removing them from the proof of Lebesgue's theorem on the existence of uncountably many Borel classes or that of Hausdorff and Alexandroff on the power of measurable (B) sets.

8 Baire functions

9 Baire's problem and the auxilliary problem of de la Vallée Poussin

10 Souslin's (A)-sets

Translator's note

Kuratowski credits his Theorem III and its Corollaries to Gerhard Hessenberg [Hes08, p.127], who was a collaborator of Ernst Zermelo.

Zermelo had in fact used this Theorem to prove his own, (see the last paragraph of [vH67, p.184]). However, his version of the proof only works in the case where the function G removes a *single* element.

I tried to translate Hessenberg's paper with the help of a German mathematician, looking for this proof, but we found the notation impenetrable.

The Theorem became known (of course many decades later) as the Bourbaki–Witt Theorem [Bou49, Wit51]. 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 than the present paper, namely by John von Neumann [vN23] and [vN28, Section 3].

References

[Bai99] René Louis Baire. *Sur les Fonctions de Variables Réelles.* PhD thesis, Paris, 1899.

[Ber08] Felix Bernstein. Über eine Eigenschaft der Mengen von Punkten auf der Geraden. *Berichte über die Verhandlungen der Königlich Sächsischen Gesellschaft der Wissenschaften zu Leipzig, Mathematisch-Physische Klasse*, 60:119–159, 1908.

[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/`.

[Bro10] L.E.J. Brouwer. On the structure of perfect sets of points. *Procedings of the Royal Academy of Amsterdam*, 12:785–794, 1910.

[Cam78] Paul Campbell. The origin of Zorn's lemma. *Historia Mathematica*, 5:77–89, 1978.

[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 [Ded60].

[Ded60] J. W. Richard Dedekind. The nature and meaning of numbers. In *Essays on the Theory of Numbers*. 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/`.

[Fré06] Maurice Fréchet. Sur quelques points du calcul fonctionnel. *Rendiconti del Circolo Matematico di Palermo*, 22:1–72, 1906.

[Fré10] Maurice Fréchet. Les ensembles abstraits et le calcul fonctionnel. *Rendiconti del Circolo Matematico di Palermo*, 30:1–26, 1910.

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

[Hau06] Felix Hausdorff. Untersuchungen über Ordnungstypen. *Berichte über die Verhandlungen der Math.–Phys. Klasse der Königlich Sächsischen Gesellschaft der Wissenschaften zu Leipzig*, 1906. vol 58 pp 106–169 and vol 59 pp 84–159.

[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.

[Jan10] Zygmunt Janiszewski. Sur la géometrie de lignes cantoriennes. *Comptes Rendus Hebdomadaires des Séances de l'Académie des Sciences de Paris*, 151:198–201, 1910.

[Jan12] Zygmunt Janiszewski. Sur les continus irréductibles entre deux points. *Journal de l'Ecole Polytechnique*, II(16):85–170, 1912. PhD Thesis.

[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.

[Mah11] Paul Mahlo. Über lineare transfinite Mengen. *Berichte über die Verhandlungen der Königlich Sächsischen Gesellschaft der Wissenschaften zu Leipzig, Mathematisch-Physische Klasse*, 63:187–225, 1911.

[Maz10] Étienne (Stefan) Mazurkiewicz. Sur la théorie des ensembles. *Comptes Rendus Hebdomadaires des Séances de l'Académie des Sciences de Paris*, pages 296–8, 1910.

[Maz19] Étienne (Stefan) Mazurkiewicz. Nowy dowód twierdzenia o istnieniu kontinuów nieprzywiedlnych — novelle démonstration du théorème sur l'existence de continus irréducibles. *Bull Acad Polon Sci Ser Math Astron Phys*, pages 44–46, 1919.

[Sie18] Wacław Sierpiński. L'axiome de M. Zermelo et son rôle dans la théorie des ensembles et l'analyse. *Bulletin international de l'Académie des sciences de Cracovie, Classe des sciences mathématiques et naturelles, Série A*, pages 97–152, 1918.

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

[Sie21] Wacław Sierpiński. Les exemples effectifs et l'axiome du choix. *Fundamenta Mathematicae*, 2(2):112–118, 1921.

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

[vN23] John von Neumann. Zur Einführung der transfiniten Zahlen. *Acta litterarum ac scientiarum Regiae Universitatis Hungaricae Franscisco-Josephinae, Sectio scientiarum mathematicarum*, 1:199–208, 1923. English translation in [vH67], pages 346–354.

[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.

[You03] William Henry Young. Note on the analysis of linear sets of points. *Quarterly Journal of Pure and Applied Mathematics*, 85, 1903.

[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.

[Zor10] Ludovic Zoretti. Sur la notion de ligne. *Comptes Rendus Hebdomadaires des Séances de l'Académie des Sciences de Paris*, 151:201–3, 1910.

[Zor35] Max Zorn. A remark on transfinite algebra. *Bull. Amer. Math. Soc.*, 41(10), 1935.