On the Problem of Well-Ordering

Friedrich Hartogs
translated by Paul Taylor

We present a proof that any set can be well ordered. Ours is different from the two that Zermelo has given [Zer04, Zer08a], because we rely on a different principle, called Comparability of Sets, instead of his Axiom of Choice. It says that, for any two given infinite sets, one must be bijective with a subset of the other. In fact, we only use this assumption in the case where one of the two sets is well ordered. Therefore the Well Ordering Principle, Comparability of Sets and the Axiom of Choice are all equivalent.

Without using any of these assumptions, our arguments still show that no set has cardinality greater than that of every well ordered set. So far as I am aware, this theorem has not previously been proved rigorously without using Choice.

The following development is based on Zermelo's axioms [Zer08b], of course omitting the Axiom of Choice (Axiom VI). It is difficult to be sure that the ideas that we use here are all independent of Choice, since Zermelo's system does not yet include a treatment of ordered and well ordered sets, but we consider this in the appendix.

We first list the requirements for the proof, maybe at the cost of brevity. Then we may see that we are using as few concepts and propositions of set theory as possible, and can check their independence from Choice. In fact, we could use ordinal numbers to make some simplifications.

Concepts:

- Element and set, equivalence (bijection) of two sets, larger and smaller cardinality. That these notions are independent of Choice follows directly from Zermelo's paper.

- Ordered set, isomorphism of two ordered sets, well- ordered set, initial segment of a well ordered set.

Lemma 1 *If two ordered sets are isomorphic to a third, they are also isomorphic to each other [Can95, p497].*

Lemma 2 *Each subset of a well ordered set is itself well ordered [Can97, p209, Prop C].*

Lemma 3 *Any set that is bijective with a well ordered set is again well ordered [Can97, Prop D].*

Lemma 4 *If, for each element of an ordered set L, its predecessors form a well ordered set, then L is itself well ordered.*

Proof: Let $a \in U \subset L$. If a is not already the first element of U, consider the set of elements of U that precede it. This is well ordered, being a subset of the well ordered set of predecessors of a in L. Its first element is also the first element of U.

Lemma 5 *A well ordered set is not isomorphic to a subset of any of its proper initial segments [Can97, p211, Prop B].*

Lemma 6 *Two well ordered sets are either isomorphic to one another or one of them is isomorphic to an initial segment of the other [Can97, p215, Prop N].*

Finally, we will frequently use the following properties of sets themselves. See the appendix for the proofs using Zermelo's theory.

Lemma 7 *For each set M there is a set $\mathfrak{m}^0$ that contains as elements all ordered*[1] *sets whose elements are identical with those of M.*

This does not make any claim that any such ordered sets have to exist, *i.e.* that any given set M can be ordered. Rather, if M cannot be ordered, $\mathfrak{m}^0$ simply means the empty set. Hence, by applying the statement to all subsets of a given set M and using Zermelo's axioms IV (Power Set) and V (Union), we obtain

Lemma 8 *For each set M there is a set $\mathfrak{m}_0$ whose elements are all ordered sets whose elements are also elements of M.*

Now, using Axiom III (Separation), those elements of $\mathfrak{m}_0$ that are *well* ordered sets form a subset $\mathfrak{n} \subset \mathfrak{m}_0$, which is also a set. Hence we have the narrower version of the previous statement that will be used later:

Lemma 9 *For every set M there is a set $\mathfrak{n}$ whose members are the well ordered sets whose elements belong to M.*

Of course $\mathfrak{n}$ is only nonempty, if M itself is.

After these preliminary remarks, we can proceed to the main proof.

Given any nonempty set M, whether finite or infinite, we may consider all possible well ordered sets[2] G, H, ... whose elements belong to M. In this we include the empty set, which has no elements. By the previous remarks, there is a nonempty set $\mathfrak{n}$ consisting of the well ordered subsets G, H, ..., and only these.

Since it is definite[3] for each of the sets G, H, ..., whether G is isomorphic to H or not, we may use Zermelo's Axiom III (Separation) to collect those that are isomorphic to G into a set $\mathfrak{g}$. Similarly the ones isomorphic to H form a set $\mathfrak{h}$ and so on.

For convenience we refer to the resulting subsets $\mathfrak{g}$, $\mathfrak{h}$, ... of $\mathfrak{n}$ as [equivalence or isomorphism] "classes". One of these classes consists only of the empty set, so we call it the null class.

From Lemma 1 it follows immediately that any two classes that have some common member are actually identical. Consequently, each member of $\mathfrak{n}$ (*i.e.* each of the well ordered sets G, H, ...) belongs to one and only one class.

We now consider the set $\mathfrak{L}$ of classes that are completely different from one another. This exists, using Axioms IV (Powerset) and III (Separation), as a subset of the power set $\mathcal{P}(\mathfrak{n})$ of $\mathfrak{n}$, since for each subset of $\mathfrak{n}$ it is definite whether it is one of our classes or not.

[1] The difference between this lemma and the next is that $\mathfrak{m}^0$ here is the set of all order relations $(<)$ on M itself, whereas $\mathfrak{m}_0$ consists of all $(U, <)$ where $U \subset M$ and $(<)$ is an order relation on U. Hartogs seems to be unsure show how such $(<)$ are to be encoded, but provides a way in the appendix .

[2] These are not just subsets of M but all of the subsets *equipped with* all possible well-orderings.

[3] Saying that "it is definite whether some property holds" means that a certain predicate is being defined in higher order logic. However, the notion of predicate was absent from Zermelo's formulation, in particular the Axiom of Separation. For that reason it seems appropriate to keep the awkward language.

Let $\mathfrak{g}$ and $\mathfrak{h}$ be two different classes, G an element of $\mathfrak{g}$ and H an element of $\mathfrak{h}$. Then the well ordered sets G and H are not isomorphic to each other; so by Lemma 6 either G is a isomorphic to a proper initial segment of H ($G < H$) or H is to one of G ($H < G$).

If $G < H$ and G' is any other element of the class $\mathfrak{g}$ to which G belongs and similarly H' is one of $\mathfrak{h}$ then we also have $G' < H'$, and *vice versa*. That is, *either every* pair G', H' of elements of classes $\mathfrak{g}$ and $\mathfrak{h}$ satisfy $G' < H'$, *or else every* pair has $H' < G'$.

In the first case we abbreviate this to $\mathfrak{g} \prec \mathfrak{h}$ and in the second $\mathfrak{h} \prec \mathfrak{g}$.

If $\mathfrak{g}$, $\mathfrak{h}$ and $\mathfrak{i}$ are three different classes with $\mathfrak{g} \prec \mathfrak{h}$ and $\mathfrak{h} \prec \mathfrak{i}$, then also $\mathfrak{g} \prec \mathfrak{i}$. Thus we have defined an ordering on the set $\mathfrak{L}$.

We now show that $\mathfrak{L}$ is *well* ordered.

Let $G \in \mathfrak{g} \in \mathfrak{L}$ and $\mathfrak{h}$ be any predecessor of $\mathfrak{g}$ ($\mathfrak{h} \prec \mathfrak{g}$), with $H \in \mathfrak{h}$. Then H is isomorphic to an initial segment H' of G and H' is then also an element of $\mathfrak{h}$. So among the members of the class $\mathfrak{h}$, there is a certain initial segment H' of G, and only one because the members of $\mathfrak{h}$ are all isomorphic sets.

Conversely, every initial segment H of G is a member of a certain unique class $\mathfrak{h} < \mathfrak{g}$.

The classes $\mathfrak{h} \prec \mathfrak{g}$ and the initial segments $H \subset G$ therefore correspond to one another unambiguously, preserving order.

Hence, by Lemma 3, the classes $\mathfrak{h} \prec \mathfrak{g}$ (*i.e.* the members of $\mathfrak{L}$ that precede $\mathfrak{g}$), form a well ordered set. But since $\mathfrak{g}$ was an arbitrary element of $\mathfrak{L}$, then by Lemma 4, $\mathfrak{L}$ is itself well ordered. (The initial element of $\mathfrak{L}$ is the null class.)

The following also applies: *Each of the well ordered sets G, H, ... is isomorphic to an initial segment of $\mathfrak{L}$.*

More precisely: if the well ordered set G belongs to the class $\mathfrak{g}$ then G is isomorphic to the initial segment of $\mathfrak{L}$ consisting of the predecessors of $\mathfrak{g}$ in $\mathfrak{L}$.

This is because, by the previous comments, there is a one-one order-preserving relationship between the classes $\mathfrak{h} \prec \mathfrak{g}$ (*i.e.* the elements of $\mathfrak{L}$ preceding $\mathfrak{g}$) and the initial segments $H \subset G$. This remains the case if instead of the initial segments of G we use the elements of G.

From this follows immediately: *$\mathfrak{L}$ can neither be bijective with the set M itself nor to any of its subsets.*

Because if $\mathfrak{L}$ were bijective with some subset U of M (where U could be the set M itself), *i.e.* if there were a mutually unambiguous assignment between the elements of the set $\mathfrak{L}$ and those of the set U, then U would itself be well ordered and isomorphic to $\mathfrak{L}$.

On the other hand, however, U is a well ordered subset of M and therefore isomorphic to an initial segment of $\mathfrak{L}$. This is a contradiction.

It is thus established that between the sets $\mathfrak{L}$ and M with cardinalities λ and μ, neither of the relationships $\lambda = \mu$ nor $\lambda < \mu$ can hold.

So if M is an arbitrary set, there always exists a well ordered set $\mathfrak{L}$ with the property that $\lambda < \mu$ does not hold.

Hence, without using the axioms of Choice or Comparability, we have proved there cannot be a quantity whose cardinality is greater than that of any well ordered set.

Now assume the Axiom of Comparability. Then it follows that, for the sets M and $\mathfrak{L}$ above, we must have $\mu < \lambda$. That is, there is a mutually unique relationship between the elements of M and those of a certain subset $\mathfrak{L}_0$, of $\mathfrak{L}$.

But since $\mathfrak{L}$ is well ordered, this relationship makes M well ordered.

Appendix

I am grateful to Mr. Hessenberg for telling me that Lemmas 7–9 can be proved from Zermelo's axioms without using Choice.

Cantor's definition of an ordered set may be replaced with an equivalent one using Zermelo's axioms, by using a subset $\mathfrak{p}$ of the power set $\mathcal{P}(M)$ that has the following three properties:

1. If R and S are two different elements of $\mathfrak{p}$, then either S is a subset of R or R is a subset of S.
2. If x and y are two different elements of M, there is some element R of $\mathfrak{p}$ that contains exactly one of x or y as an element.
3. The union set $\bigcup \mathfrak{p}'$ of any subset $\mathfrak{p}'$ of $\mathfrak{p}$ is an element of $\mathfrak{p}$.

This definition is equivalent to Cantor's for the following reason: Given a set M that is ordered in the Cantorian sense, then its collection of residues has these three properties.

Conversely, if $\mathfrak{p}$ is a set of this form and we define $x \prec y$ if there is an element R of $\mathfrak{p}$ that contains y but not x, then we have an ordering in the Cantorian sense, namely the set of the residues of the latter is identical to $\mathfrak{p}$.

Since according to this definition, every order relation on M is a subset of $\mathcal{P}(M)$ and it is definite whether any subset of $\mathcal{P}(M)$ is an order or not. Hence the totality $\mathfrak{O}$ of all orders of M is a subset of $\mathcal{P}\mathcal{P}M$, and at any rate a set, which proves Lemma 7.

In the proof of Lemma 9 from Lemma 8, we make use of the fact that, among the orderings of M, the well-orderings are distinguished by a definite criterion in the sense of Zermelo.

This criterion is that the set $\mathfrak{p}$ defined above represents a well-ordering of M if the union $\bigcup \mathfrak{p}'$ of any subset $\mathfrak{p}'$ of $\mathfrak{p}$ an element of $\mathfrak{p}'$, not just of $\mathfrak{p}$.

Note that Lemma 8 and 9 can also be proved without using Lemma 7. The above property 2 of the set $\mathfrak{p}$ can be resolved into to two parts:

2a. The union $\bigcup \mathfrak{p}$ of $\mathfrak{p}$ is $\mathcal{P}M$.

2b. If x and y are two different elements of $\bigcup \mathfrak{p}$, then there is some element R of $\mathfrak{p}$ that contains exactly one of them as an element.

If we continue with property 2a, then in general $\mathfrak{p}$ no longer defines an order on M, but rather an order on a subset U of it.

According to this, $\mathfrak{m}_0$ is also a subset of $\mathcal{P}\mathcal{P}M$.

Finally, if we replace 3 by the further requirement that the union $\bigcup \mathfrak{p}'$ of each subset $\mathfrak{p}'$ of $\mathfrak{p}$ is an element of $\mathfrak{p}'$, then $\mathfrak{n}$ is also expressed directly as a subset of $\mathcal{P}\mathcal{P}M$.

Details of this theory of ordered and well ordered sets may be found in Hessenberg's book [Hes06, §28] and paper [Hes09] and in Zermelo's second proof of well-ordering [Zer08a].

Munich, July 1914.

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.

[Can95] Georg Cantor. Beiträge zur Begründung der transfiniten Mengenlehre (Erste Artikel). *Mathematische Annalen*, 46:481–512, 1895. Reprinted in [Can32]; English translation in [Can15].

[Can97] Georg Cantor. Beiträge zur Begründung der transfiniten Mengenlehre (Zweiter Artikel). *Mathematische Annalen*, 49:207–246, 1897. Reprinted in [Can32]; English translation in [Can15].

[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.

[Har15] Friedrich Hartogs. Über das Problem der Wohlordnung. *Mathematische Annalen*, 76:438–443, December 1915. `zenodo.org/record/1428298`.

[Hes06] Gerhard Hessenberg. *Grundbegriffe der Mengenlehre*, volume 1. Abhandlungen der Friesschen Schule, 1906. `openlibrary.org/books/OL20528826M`.

[Hes09] Gerhard Hessenberg. Mengenlehre. In Felix Auerbach and Rudolf Ernst Rothe, editors, *Taschenbuch für Mathematiker und Physiker*, pages 60–63. Teubner, Leipzig, 1909. `ia802704.us.archive.org/33/items/taschenbuch-fur-mathematiker-und-physiker`.

[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.

[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.

[Zer10] Ernst Zermelo. *Collected Works, volume I.* Number 21 in Schriften der Mathematisch-naturwissenschaftlichen Klasse. Springer-Verlag, 2010. edited by Heinz-Dieter Ebbinghaus, Craig Fraser and Akihiro Kanamori, including parallel English translations.

Friedrich Moritz Hartogs was a German Jew, born on 20 May 1874 in Brussels but brought up in Frankfurt-am-Main. He was a student in Berlin and became a Privatdozent and then a full Professor in Munich. He was fired by the Nazis in 1935 and took his own life on 18 August 1943. Besides set theory, he also studied functions of several complex variables.

I have made use of different typefaces to clarify the ranks of the (sub)sets:

- the given set is M;
- $x, y \in M$;
- $G, H \subset M$, but equipped with well-orderings;
- $\mathfrak{g}, \mathfrak{h}, \mathfrak{i}, \mathfrak{m}, \mathfrak{n}, \mathfrak{p}, \mathfrak{p}' \subset \mathcal{P}(M)$;
- $\mathfrak{L} \subset \mathcal{P}\mathcal{P}(M)$;
- λ, μ are the cardinalities of $\mathfrak{L}, M$.

In the construction above, $\mathfrak{L}$ is the smallest *ordinal* whose cardinality is strictly bigger than that of the given set M, since any smaller ordinal is some $\mathfrak{g} \in \mathfrak{L}$, whose predecessors are bijective with some $G \subset M$. By putting a well-ordering on any other set K, it follows that $\mathfrak{L}$ is the also successor (next biggest) *cardinal* of that of M.

This result may be used as a way of saying "when to stop" in transfinite recursion over the ordinals. Such recursion was justified by [vN28, Thm 3]. It provides a function $r : \mathfrak{L} \to M$, which cannot be injective, so there must be some $\mathfrak{g} < \mathfrak{h} \in \mathfrak{L}$ with $r(\mathfrak{g}) = r(\mathfrak{h})$. Since r is typically defined to preserve order, it follows that $r(\mathfrak{g}) \leq r(\mathfrak{g}+1) = s(r(\mathfrak{g})) \leq r(\mathfrak{h})$, so $r(\mathfrak{g}) \in M$ is a fixed point of the function $s : M \to M$ from which the recursion is defined.

Hartogs' construction does not actually depend on the precise formulation of the definition of ordinal, only on the lemmas about them that he lists. In fact Cesare Burali Forti [BF97] had proved his famous "paradox" using a mistaken definition.

Work in the 1990s showed that there are several constructive notions of ordinal and Hartogs' construction is still valid for them. Unfortunately, it cannot be *applied* constructively, because in the remark above we actually only have $\neg\neg r(\mathfrak{g}) = r(\mathfrak{h})$.

The representation of orders that Hartogs attributes to Hessenberg is nowadays known as the Alexadrov Topology. It is also the poset version of what is now called the Yoneda embedding of a category.

Axiom 3 is the (now) familiar union axiom for open subsets in a topology. Axiom 1 makes the topology linearly ordered, so the binary intersection and union are the smaller and larger of the pair, so the infinitary union is in fact directed. Axiom 2 makes the space T_1 and open in itself.

The orientation is that open subsets are upper, whereas predecessors and initial segments are lower.

In the representation of *well* orderings, even the directed unions become trivial, capturing the characterisation of well-orderings that they have no infinite descending sequences. However, this characterisation relies on Dependent Choice, which may be at odds with Hartogs' claim that is not using (full) Choice.