Chapter I. Some choiceless results

The topic of this course is Combinatorial set theory, so even though we will study additional axioms, we will not emphasize forcing or inner model-theoretic techniques. Similarly, we will study some results of a descriptive set theoretic nature, but will not delve into the fine definability issues that descriptive set theory involves. With a few exceptions that will be clearly indicated, we will assume the axiom of choice throughout (and I will assume basic knowledge of axiomatic set theory, cardinals and ordinals), but we begin by looking at some results that do not require the axiom of choice.

1 Cantor’s theorem

Theorem 1 If \( f : X \to \mathcal{P}(X) \) then \( f \) is not surjective.

Proof: Let \( A = \{ y \in X : y \notin f(y) \} \). Then \( A \notin \text{ran}(f) \). \( \square \)

Theorem 2 If \( f : \mathcal{P}(X) \to X \) then \( f \) is not injective.

Proof: Let \( A = \{ y \in X : \exists Z (y = f(Z) \notin Z) \} \) and set \( a = f(A) \), so \( a \in A \), so there is some \( Z \neq A \) with \( f(Z) = f(A) \). \( \square \)

Note that in Theorem 1 we explicitly (i.e., definably) found a set not in the range of \( f \). In Theorem 2, we found a set \( A \) for which there is a set \( Z \) with \((A, Z)\) witnessing a failure of injectivity, but we did not actually define such a set \( Z \). I do not know whether this can be done; we will see later a different argument in which such a pair is defined.

2 The Knaster-Tarski theorem

This section was originally a handout I wrote for a set theory course I taught at Caltech. Recall:

Definition 3 A complete lattice is a poset \( \mathcal{P} = (P, \leq) \) such that any subset of \( P \) has an inf and a sup.

Theorem 4 (Knaster-Tarski) If \( (\mathcal{L}, \leq) \) is a complete lattice and \( f : \mathcal{L} \to \mathcal{L} \) is order-preserving (also called monotone or isotone), then \( f \) has a fixed point. In fact, the set of fixed points of \( f \) is a complete lattice.

Proof: Let \( M = \{ x \in \mathcal{L} : x \leq f(x) \} \). Since \( \mathcal{L} \) has a minimum, \( M \) is nonempty. Since \( f \) is order-preserving, \( f''M \subseteq M \). Let \( x \in M \) and \( x_1 = \text{sup} M \). Then \( x \leq x_1 \), so \( f(x) \leq f(x_1) \).
but also \( x \leq f(x) \), so \( f(x_1) \) is an upper bound for \( M \). Hence, \( x_1 \leq f(x_1) \), so \( x_1 \in M \). But then \( f(x_1) \in M \), so \( f(x_1) \leq x_1 \), and it follows that \( f(x_1) = x_1 \).

Let \( S \) be the set of fixed points of \( L \). Clearly, \( x_1 = \max S \). The same argument as above shows that if \( x_0 = \inf \{ x \in L : f(x) \leq x \} \) then \( x_0 = \min S \).

Let \( \emptyset \neq X \subseteq L \). Let \( y = \sup X \). Then \( f(y) \) is also an upper bound for \( X \), so \( y \leq f(y) \).

Now consider the set \( S = \{ x \in L : y \leq x \} \). With the order inherited from \( L \), \( S \) is itself a complete lattice. Let \( f' = f \upharpoonright S \). Then \( f' \) is an order-preserving map of \( S \) to itself. It follows that there is a least fixed point \( y_1 \) of \( f' \). Clearly, \( y_1 \in S \) and in fact \( y_1 = \sup^S X \), where \( \sup^S \) means that the sup is computed from the point of view of \( S \) (but \( y < y_1 \) is possible). Similarly one shows that \( y_0 = \inf^S X \) exists and therefore \((S, \leq)\) is a complete lattice—although, in general, it is not a complete sublattice of \( L \). □

Notice that to show that \( S \neq \emptyset \) we only used that \( L \) has a minimum and every subset of \( L \) has a sup. Let us call such partially ordered sets complete join semilattices. The reason why we stated the existence of fixed points for complete lattices as opposed to complete join semilattices is the following observation:

**Lemma 5** Any complete join semilattice is a complete lattice.

**Proof:** Let \( (L, \leq) \) be a complete join semilattice and set \( 0 = \min L \). For \( a \in L \) let \( [0, a] = \{ x \in L : x \leq a \} \). Let \( \emptyset \neq A \subseteq L \). Then \( b = \sup \bigcap_{a \in A} [0, a] \) exists, and it is straightforward to verify that \( b = \inf A \). □

**Definition 6** A lattice is a nonempty poset (with a max and a min) such that any finite subset has a max and a min.

The following argument uses the axiom of choice.

**Theorem 7** (Davis) If \( (L, \leq) \) is a lattice such that any order-preserving function \( f : L \to L \) has a fixed point, then \( L \) is a complete lattice.

**Proof:** Assume that \( L \) is not complete. The argument proceeds by building a gap: Two sequences in \( L \), \( (a_\xi : \xi < \alpha) \) and \( (b_\nu : \nu < \beta) \), such that

- \( a_\xi < b_\nu \) for all \( \xi < \alpha, \nu < \beta \).
- \( (a_\xi)_{\xi < \alpha} \) is strictly increasing.
- \( (b_\nu)_{\nu < \beta} \) is strictly decreasing.
- There is no \( c \in L \) such that \( a_\xi < c < b_\nu \) for all \( \xi < \alpha, \nu < \beta \).

This is easily accomplished by an argument involving Zorn’s lemma. We first claim that there exists a decreasing sequence \( C \subseteq L \) that has no inf. We show this by arguing that,
otherwise, every subset $S \subseteq \mathcal{L}$ has a sup, so $\mathcal{L}$ is a complete join semilattice, contradicting Lemma 5: The proof is similar to the proof of Lemma 5. Let $U$ be the set of upper bounds of $S$, so $U \neq \emptyset$ since $\max \mathcal{L} \in S$. Let $C \subseteq U$ be a maximal decreasing sequence. Then $c = \inf C$ exists, by our assumption, and it is easy to see that $c = \sup S$.

We have shown that we can find a strictly decreasing sequence $(b_\nu)_{\nu < \beta}$ without an inf. Now consider a maximal increasing sequence $(a_\xi)_{\xi < \alpha}$ of lower bounds of $\{b_\nu : \nu < \beta\}$. Then $(a_\xi)_{\xi < \alpha}$ and $(b_\nu)_{\nu < \beta}$ are as required.

Now we define a function $f : \mathcal{L} \to \mathcal{L}$ and proceed to verify that $f$ is order-preserving and has no fixed points. Let $x \in \mathcal{L}$. Assume first that $x$ is a lower bound of $(b_\nu)_{\nu < \beta}$. Then there must be a least $\xi < \alpha$ such that $a_\xi \not\leq x$. Let $f(x) = a_\xi$ in this case.

If $x$ is not a lower bound of $(b_\nu)_{\nu < \beta}$, let $f(x) = b_\nu$, where $\nu < \beta$ is least such that $x \not\leq b_\nu$.

By definition, either $x \not\leq f(x)$ or $f(x) \not\leq x$, so $f$ has no fixed points. We conclude by showing that $f$ is order-preserving. For this, suppose that $x \leq y$. If $y$ is a lower bound of $(b_\nu)_{\nu < \beta}$, let $f(y) = a_\xi$. Then $a_\xi \not\leq x$, or else $a_\xi \leq y$. Hence, $f(x) = a_\xi'$ for some $\xi' \leq \xi$ (since $x$ is also a lower bound of $(b_\nu)_{\nu < \beta}$), so $f(x) \leq f(y)$.

We can thus assume that $y$ is not a lower bound of $(b_\nu)_{\nu < \beta}$. Let $f(y) = b_{\nu_0}$. If $x$ is not a lower bound either, then (since $x \leq y$), $f(x) = b_{\nu_1}$ for some $\nu_1 \geq \nu_0$, so $f(x) \leq f(y)$.

Finally, if $x$ is a lower bound, then $f(x) = a_\xi$ for some $\xi$, but then $f(x) < f(y)$.

2.1 Historical remarks

Theorem 4 was shown by Knaster and Tarski in 1927 for a particular case, see [8], namely, for any $S$, any $\subseteq$-increasing $\pi : \mathcal{P}(S) \to \mathcal{P}(S)$ has a fixed point. The Schröder-Bernstein theorem discussed below (see Theorem 8) and applications to topology are mentioned in [8]. The general version [15] appeared in 1955 but Tarski mentions it in lectures from 1939. Around the time [15] was being prepared for publication, Tarski asked about the converse, and soon after Anne Davis proved Theorem 7. Davis’s paper [2] appears in the same journal as Tarski’s, immediately following his paper.


2.2 Applications

**Theorem 8 (Schröder-Bernstein)** If there are injections $f : A \to B$ and $g : B \to A$ then there is a bijection $h : A \to B$.

**Proof:** Consider $\pi : \mathcal{P}(A) \to \mathcal{P}(A)$ given by $\pi(X) = A \setminus g''(B \setminus f''X)$. Clearly, $\pi$ is $\subseteq$-preserving. Let $X_0$ be a fixed point of $\pi$. Then define $h = (f \upharpoonright X_0) \cup (g^{-1} \upharpoonright A \setminus X_0)$. □

The “usual” proof of Theorem 8 proceeds by explicitly building a fixed point $X_0$ by a back-and-forth argument, and then defining $h$ as above. For example, Kunen [10] sets
A_0 = A, B_0 = B, A_{n+1} = g''B_n, B_{n+1} = f''A_n, and \(A_\omega = \bigcap_n A_n\), and considers \(X_0 = A_\omega \cup \bigcup_n (A_{2n} \setminus A_{2n+1})\). This corresponds to the largest fixed point of \(\pi\). Other textbooks construct the smallest fixed point (by removing \(A_\omega\) from \(X_0\)).

Of course, particular versions of Theorem 4 can be proved by building fixed points of the relevant maps by assuming some form of continuity and then iterating the maps. For example, one has:

**Theorem 9 (Tarski-Kantorovitch)** Let \(\mathbb{P} = (P, \leq)\) be a poset and let \(f : P \to P\) be sequentially continuous, i.e., if \((c_n)_{n<\omega}\) has a supremum \(c\), then \(f(c) = \sup_n f(c_n)\). Assume that there is \(b \in P\) such that \(b \leq f(b)\) and any countable chain in \(\{x : x \geq b\}\) has a supremum. Let \(b' = \sup_n f^n(b)\), where \(f^n\) is the \(n\)-th iterate of \(f\). Then \(b'\) is a fixed point of \(f\), and it is the least fixed point of \(f\) in \(\{x : x \geq b\}\).

**Proof:** The obvious argument works. Just notice that \(f\) is order preserving since if \(y \geq x\) then \(y = \sup \{x, y\}\), so \(f(y) = \sup \{f(x), f(y)\}\). \(\Box\)

Other applications of Theorem 4 come from computer science (in denotational semantics), but here, again, one tends to be interested in “explicit” constructions of fixed points. Trans-finite recursion allows us to do this: Consider an order-preserving \(f : \mathcal{L} \to \mathcal{L}\) for \(\mathcal{L}\) a complete lattice. Let \(f^\alpha\) be the \(\alpha\)-th iterate of \(f\), so \(f^0\) is the identity, \(f^{\alpha+1} = f(f^\alpha)\) and for \(a \in \mathcal{L}\), \(f^\lambda(a)\), for \(\lambda\) limit, is the sup of \(\{f^\alpha(a) : \alpha < \lambda\}\). For any \(a\), the sequence \((f^\alpha(a) : a \in \text{ORD})\) is (not necessarily strictly) increasing, so it eventually stabilizes, and necessarily this stable value is a fixed point of \(f\). For \(a = 0\) this gives the least fixed point of \(f\). If instead we define \(f^\lambda\), for \(\lambda\) limit, as the infimum, then for \(a = 1\) the sequence provides us with the largest fixed point.

An elaboration of the proof of Theorem 8 shows the following:

**Theorem 10 (Tarski)** Let \(\mathbb{B} = (B, \leq)\) be a complete Boolean algebra. Let \(a, b \in B\) and let \(f : [0, a] \to B\) and \(g : [0, b] \to B\) be increasing. Then there are \(a', b' \in A\) such that \(f(a - a') = b'\) and \(g(b - b') = a'\). \(\Box\)

This result has nice applications in the theory of complete Boolean algebras. We mention one from [15]: Say that two elements \(a\) and \(b\) of a Boolean algebra \(\mathbb{B}\) are homogeneous, \(a \approx b\), iff \(([0, a], \leq)\) and \(([0, b], \leq)\) are isomorphic.

**Theorem 11 (Tarski)** Let \(\mathbb{B} = (B, \leq)\) be a complete Boolean algebra. Let \(a, b, c, a', c' \in B\) be such that \(a \leq b \leq c, a' \leq c', a \approx a'\) and \(c \approx c'\). Then there is \(b'\) such that \(a' \leq b' \leq c'\) and \(b \approx b'\). \(\Box\)

**Definition 12** A derivative algebra is a structure \(A = (A, \leq, d)\) where \((A, \leq)\) is a Boolean algebra and \(d : A \to A\) satisfies the following equations:

- \(d(0) = 0\).
• $d dx \leq x + dx$ for all $x \in A$.

• $d(x + y) = (dx) + (dy)$ for all $x, y \in A$.

(Where, of course, $x + y = \sup\{x, y\}$ and $0 = \inf A$.)

There is a standard example of derivative algebras, and a characterization theorem:

• Let $X$ be a topological space. Then $P(X)$ is a derivative algebra with $\subseteq$ as $\leq$ and $dX$ the standard derivative of the set $X$, i.e., the set of all limit points of $X$.

A modal logic is obtained by adding to propositional logic a modal operator $\Box$ (read it is necessary that), so if $\varphi$ is a formula, so is $\Box \varphi$. The modal logic $wK4$ is defined by the axioms of propositional logic and the two modal axioms $\Box(p \rightarrow q) \rightarrow (\Box p \rightarrow \Box q)$ and $(p \land \Box p) \rightarrow \Box \Box p$, with inference rule Modus Ponens and $p \vdash \Box p$.

In the same sense that Boolean algebras correspond to propositional logic, derivative algebras correspond to the modal logic $wK4$. (This can be made more precise in the study of Heyting algebras, that correspond to intuitionistic logic.)

Let $A = (A, \leq, d)$ be a derivative algebra. An element $a \in A$ is closed iff $da \leq a$; it is dense in itself iff $da \geq a$; it is perfect iff $da = a$; and it is scattered iff there is no dense in itself element $x$ such that $0 \neq x \leq a$.

The following generalizes the Cantor-Bendixon theorem:

**Theorem 13 (Tarski)** Let $A = (A, \leq, d)$ be a derivative algebra such that $(A, \leq)$ is complete. Then every closed $a \in A$ admits a decomposition $a = b + c$, $b \times c = 0$, with $b$ perfect and $c$ scattered.

**Proof:** Let $b = \sup\{x: a \times dx \geq x\}$ and let $c = a - b$. Then $a = b + c$ and $b \times c = 0$. Now let $d_a : A \rightarrow A$ be given by $d_a(x) = a \times dx$. Clearly, $d_a$ is order-preserving, so (by the proof of Theorem 4) $b$ is the largest fixed point of $d_a$: $b = d_a b = a \times db$. Then $b \leq a$ and $db \leq da$. But $a$ is closed, so $db \leq a$ and since $b = a \times db$, $b = db$, i.e., $b$ is perfect.

Let $x \leq c$ be dense in itself, so $dx \geq x$. Then $d_a x \geq x$ (since $a \geq c$), so $x \leq b$. Hence, $x = 0$. It follows that $c$ is scattered and we are done. $\square$

### 3 The dual Schröder-Bernstein theorem.

The Schröder-Bernstein Theorem 8 was proved above as a corollary of the Knaster-Tarski Theorem 4.

Another nice way of proving the result is graph theoretic. We may assume that $A$ and $B$ are disjoint and form a directed graph whose nodes are elements of $A \cup B$ and there is an edge from $a$ to $b$ iff either $a \in A$ and $b = f(a)$ or $a \in B$ and $b = g(a)$. Consider the connected
components of this graph. Each component is either a cycle of even length, or a $Z$-chain, or a $N$-chain. In each case, one can canonically find a bijection between the elements of the component in $A$ and the elements in $B$. Putting these bijections together gives the result.

Cantor’s proof of this result uses the axiom of choice (in the form: every set is in bijection with an ordinal).

There are a few additional remarks on the Schröder-Bernstein theorem worth mentioning.

The **dual Schröder-Bernstein theorem (dual S-B)** is the statement “Whenever $A,B$ are sets and there are surjections from $A$ onto $B$ and from $B$ onto $A$, then there is a bijection between $A$ and $B$.”

- This follows from the axiom of choice. In fact, AC is equivalent to: Any surjective function admits a right inverse. So the dual S-B follows from choice and the S-B theorem.

- The proofs of S-B actually show that if one has injections $f:A\to B$ and $g:B\to A$, then one has a bijection $h:A\to B$ contained in $f\cup g^{-1}$. So the argument above gives the same strengthened version of the dual S-B. Actually, over ZF, this strengthened version implies choice, as shown by Banaschewski and Moore in [1].

- If $j:x\to y$ is onto, then there is $k:\mathcal{P}(y)\to\mathcal{P}(x)$ 1-1, so if there are surjections in both directions between $A$ and $B$, then $\mathcal{P}(A)$ and $\mathcal{P}(B)$ have the same size. Of course, this is possible even if $A$ and $B$ do not.

**Open question.** (ZF) *Does the dual Schröder-Bernstein theorem imply the axiom of choice?*

- The dual S-B is not a theorem of ZF.

a) The example $A = \mathbb{R}$, $B = \mathbb{R} \times \omega_1$ in Solovay’s model (or under determinacy) shows that the dual S-B is not a theorem of ZF.

[Let me expand some on this example. There is a definable bijection $\mathbb{N} \times \mathbb{N} \leftrightarrow \mathbb{N}$, so (by Schröder-Bernstein) $(2^{\mathbb{N}}) \times (2^{\mathbb{N}}) \sim (2^{\mathbb{N}})$, where $A \sim B$ means that there is a bijection between $A$ and $B$. Clearly, $2^{\mathbb{N}} \preceq \mathbb{N}^\mathbb{N}$, where $A \preceq B$ means that there is an injection from $A$ into $B$. But also $\mathbb{N}^\mathbb{N} \preceq 2^{\mathbb{N}}$, since $\mathbb{N}^\mathbb{N} \preceq (2^{\mathbb{N}})^\mathbb{N} \preceq 2^{\mathbb{N}}$.

There is a bijection between $\mathbb{N}^\mathbb{N}$ and $(0,1) \setminus \mathbb{Q}$; for example, we can identify the sequence $(n_0,n_1,\ldots)$ with the real

$$\frac{1}{(n_0+1)} + \frac{1}{(n_1+1)} + \cdots.$$  

Finally, its is easy to check that $(0,1) \setminus \mathbb{Q} \sim \mathbb{R}$, for example, using that any countable dense subset of $(0,1)$ is isomorphic to $\mathbb{Q}$, and appealing again to the Schröder-Bernstein theorem. It follows that $\mathbb{R} \times \mathbb{R} \sim \mathbb{R}$.
There is a surjection $\mathbb{R} \to \omega_1$. For example, if the real $x$ is an irrational in $(0, 1)$, then it corresponds to a sequence $(n_0, n_1, \ldots)$. Set $n_i = 2^{m_i}(2k_i + 1) - 1$ for all $i$, let $R_x$ be the relation $\{(m_i, k_i) : i < \omega\}$ and let $A_x$ be the field of $R_x$, i.e., $A_x = \text{dom}(R_x) \cup \text{ran}(R_x)$. Then the real $x$ encodes this way a structure $(A_x, R_x)$. If this structure happens to be a well-order, it is isomorphic to a unique ordinal $\alpha$, and we say that the real $x$ codes $\alpha$. If $(A_x, R_x)$ is not a well-order, or if $x$ is not in $(0, 1) \setminus \mathbb{Q}$, we say that $x$ codes $0$. This is a surjection from $\mathbb{R}$ onto $\omega_1$.

It follows that there is a surjection from $A = \mathbb{R}$ onto $B = \mathbb{R} \times \omega_1$ (map $\mathbb{R}$ onto $\mathbb{R} \times \mathbb{R}$, and compose with the surjection just described), and there is also a surjection $\mathbb{R} \times \omega_1 \to \mathbb{R}$, namely, the projection onto the first coordinate.

So far, we have just worked under ZF. Assume now that $\omega_1 \not\subseteq \mathbb{R}$, i.e., there is no injective $\omega_1$-sequence of reals. This holds in Solovay’s model, and also under determinacy. Then there can be no bijection between $A$ and $B$.

Since Solovay’s model requires an inaccessible cardinal, this example may lead one to wonder whether in consistency strength one needs an inaccessible for an example.

b) One does not; an example has been found by Benjamin Miller, the details of this and what follows are in the note [11] I am attaching here (from September, 2008); they require a stronger background in descriptive set theory than I am assuming:

Miller’s example holds in Solovay’s model but actually one only needs the Baire property of all sets of reals (and Shelah showed this is equiconsistent with ZF). Take $A = 2^\mathbb{N}$ and $B = 2^\mathbb{N}/E_0$, where

$$xE_0y \text{ iff } \exists n \forall m \geq n \ (x(m) = y(m)).$$

Then the map which sends $x$ to $[x]_{E_0}$ is a surjection from $A$ onto $B$. To get a surjection from $B$ onto $A$, let $P$ denote the set of points of the form $(x|0)(x|1)(x|2)\ldots$, where $x$ is in $2^\mathbb{N}$. Send each equivalence class $C$ in $[P]_{E_0}/E_0$ to the unique element of $C$ in $P$, and send each point of $B \setminus [P]_{E_0}/E_0$ to $0^{\infty}$.

It remains to check that there is no bijection between $A$ and $B$. If there was, then we could use the lexicographic order on $A$ and the bijection to get a linear ordering of $B$, and since we are assuming that all sets of reals have the Baire property, this would contradict the well-known fact that there is no linear ordering of $2^\mathbb{N}/E_0$ which has the Baire property when thought of as a subset of $2^\mathbb{N} \times 2^\mathbb{N}$.

(This latter fact is proven as follows: If $\leq$ is such an ordering, then for each equivalence class $C$, either

- Comeagerly many classes are $\leq C$, or
- Comeagerly many classes are $\geq C$,

by generic ergodicity. Another application of generic ergodicity then gives that either

1. Comeagerly many classes are $\leq$ comeagerly many classes, or
2. Comeagerly many classes are $\geq$ comeagerly many classes.
The Kuratowski-Ulam theorem then implies that both 1. and 2. hold, thus there is an $E_0$-invariant comeager set $D$ such that $x \leq y$, for all $x, y \in D$, which contradicts the fact that $\leq$ is a partial order.)

c) In example a) there is an injection from $\mathbb{R}$ into $\mathbb{R} \times \omega_1$. In example b) there is an injection from $2^{\mathbb{N}}$ into $2^{\omega_1}/E_0$; one can use the $P$ above to build this injection, or appeal to Silver’s theorem. This may lead one to wonder whether in any example $A$ and $B$ are always comparable. That is not the case; and Miller’s note addresses this:

One can find a model in which there are Borel equivalence relations $E$ and $F$ on Polish $X$ and $Y$ such that there is no injection from $X/E$ to $Y/F$ or from $Y/F$ to $X/E$. For example, $E_1$ and $E_0^{\mathbb{N}}$ work; one can check that if $C$ is comeager, then there is no Baire measurable reduction of $C/E$ to $(2^{\omega})^{\omega}/F$. This gives that there is no injection of the quotient by $E$ into the quotient by $F$ in any model of $\text{ZF}+$ all sets of reals have the $\text{BP}$+ the version of uniformization for subsets of the plane given in Saharon Shelah [13]. So the consistency of this example follows from that of $\text{ZF}$.

(However, no pair of countable Borel equivalence relations have this property, although if one does not care about consistency strength and uses Lebesgue measurability instead of Baire category, then one can find $E, F$ countable Borel.)

To see that this works, one simply has to observe that, essentially, by Silver’s theorem, if $E'$ and $F'$ are any two Borel equivalence relations on Polish spaces $X'$ and $Y'$, and $E'$ has uncountably many equivalence classes, then there is a homomorphism from $X'/E'$ to $Y'/F'$ whose graph is Borel (when viewed as a subset of $X' \times Y'$); this follows for example from results of Kechris and Louveau in [9].

• Finally, the **Partition principle** is the statement that “if there is a surjection $f : A \to B$ then there is an injection $g : B \to A$.” So the axiom of choice implies the partition principle and the partition principle implies the dual S-B. It is open whether either implication is reversible. But if one thinks about example a), one sees that if the dual S-B holds then, for example, whenever $x$ is a set and $x \times x$ is equipotent to $x$ then whenever there is a surjection from $x$ onto a set $y$, then there is an injection from $y$ into $x$, so a weak version of the partition principle holds.

4 **Zermelo’s well-ordering theorem**

Here I follow the nice article Kanamori [6], (as of this writing) available [here](#).

**Theorem 14** If $F : \mathcal{P}(X) \to X$ then there is a unique well-ordering $(W, <)$, $W \subseteq X$, such that

1. $\forall x \in W \ F(\{y \in W : y < x\}) = x$, and
2. $F(W) \in W$.
Proof: There are two (essentially equivalent) ways of showing this. Assuming that the theory of ordinals has been developed, one simply notices that \( W = \{ a_\alpha : \alpha < \beta \} \) where \( a_\alpha = F(\{ a_\gamma : \gamma < \alpha \}) \) for all \( \alpha < \beta \), and \( \beta \) is largest such that the map \( \varphi = \varphi_\beta : \beta \to X \) given by \( \varphi(\alpha) = a_\alpha \) is injective. That \( \beta \) exists is a consequence of Hartogs theorem (see Theorem 23 below) that for any set there is a least ordinal that does not inject into \( X \). It follows from this that there is a least \( \beta \) such that \( \varphi_\beta \) defined as above is injective, but \( a_\beta \) has already been listed as \( a_\alpha \) for some \( \alpha < \beta \).

Another way of presenting this argument, without the need for the prior development of the theory of the ordinals is to argue as Zermelo originally did, in his 1904 paper (prior to Von Neumann’s result that each well-ordering is isomorphic to a unique ordinal). This method also has the technical advantage of not requiring the axiom of replacement. Say that \( A \subseteq X \) is an \( F \)-set iff there is a well-ordering \( < \) or \( A \) such that for all \( x \in A \), \( F(\{ y \in A : y < x \}) = x \). One then argues (by considering least counterexamples) that any two \( F \)-sets are comparable, in that one is an initial segment of the other. In particular, the well-ordering \( < \) associated to an \( F \)-set \( A \) is unique. It follows from this that \( W = \bigcup \{ F \text{-sets} \} \) is itself an \( F \)-set, and is as wanted, for if \( F(W) \notin W \), then \( W \cup \{ F(W) \} \) would also be an \( F \)-set. \( \square \)

Corollary 15 If \( P(X) \) admits a choice function, then \( X \) is well-orderable.

Proof: Let \( G : P(X) \to X \) be a choice function, so \( G(A) \in A \) for all \( \emptyset \neq A \subseteq X \). Let \( F(Y) = G(X \setminus Y) \) for \( Y \subseteq X \), and notice that if \( Y \neq X \), then \( F(Y) \notin Y \). Let now \( W \) be as in Zermelo’s theorem. Then \( W = X \) (and therefore \( X \) is well-orderable), since \( F(W) \in W \). \( \square \)

Corollary 16 If \( f : P(X) \to X \), then (definably from \( f \)) there is a pair \( (A,B) \) of distinct subsets of \( X \) such that \( f(A) = f(B) \).

Contrast this with the proof of the ‘dual’ version of Cantor’s theorem given above, Theorem 2, in which we defined from \( f \) a set \( A \) for which there is another set \( B \) with \( f(A) = f(B) \), but we failed to define some such \( B \).

Proof: Let \( (W,\prec) \) be as in Zermelo’s theorem, and set \( A = W \), \( B = \{ y \in W : y < F(W) \} \). \( \square \)

Corollary 17 Given a set \( X \), set \( W(X) \) denote the collection of well-orderable subsets of \( X \), and let \( O(X) \) denote the collection of relations \( R \subseteq X \times X \) that are well-orderings of their field. Then \( |X| < |O(X)| \) and \( |X| < |W(X)| \). \( \square \)

5 Specker’s lemma

This result comes from Specker [14]. I follow the nice paper Kanamori-Pincus [7] in the presentation of this and the following result. The Kanamori-Pincus paper, to which we will return below, also has several interesting problems, results, and historical remarks. (As of this writing) it can be found here.
Lemma 18 Assume that $m > 1$. Then $m + 1 < 2^m$.

PROOF: Assume otherwise, so there is a set $X$ with at least two elements, an $a \notin X$, and an injection $f : \mathcal{P}(X) \to X \cup \{a\}$. Necessarily, $a \in \text{ran}(f)$ (by the second version of Cantor’s theorem), so by switching two values if needed, we may assume that $f(X) = a$. Now use the argument of Zermelo’s Theorem 14 to find a well-ordering $(W, <)$, i.e., $W \subseteq X$ and for all $x \in W$, $f(\{y \in W : y < x\}) = x$. Moreover, $W$ is as large as possible.

Since $f$ is injective, it cannot be the case that $f(W) \in W$. Otherwise,

$$f(W) = f(\{y \in W : y < f(W)\}),$$

contradicting injectivity of $f$. Thus, the construction of $W$ must end because $W = X$, i.e., $X$ is well-orderable. If $X$ is finite, then $|X| + 1 < 2^{|X|}$ (as shown by, say, induction), so we must have that $X$ is infinite. But then $|X| + 1 = |X|$, and again we get a contradiction from the second version of Cantor’s theorem. □

Of course, strengthening the hypothesis one can find stronger conclusions. Specker’s paper actually shows that if $m$ is infinite then, for all finite $n$, one has $n \cdot m < 2^m$. This he deduces from the fact that if $m \geq 5$, then $2^m \not\leq m^2$. In the next section we show a recent strengthening of this result (but notice that the hypothesis we use is stronger than simply assuming that $m$ is infinite).

6 The Halbeisen-Shelah theorem

This comes from Halbeisen-Shelah [4].

Theorem 19 Assume that $\aleph_0 \leq |X|$. Then $|\mathcal{P}(X)| \not\leq |\text{Seq}(X)|$, where $\text{Seq}(X)$ denotes the set of finite sequences of elements of $X$.

PROOF: Assume otherwise, and let $G : \mathcal{P}(X) \to \text{Seq}(X)$ be injective. We use $G$ to define a function $F$ from the set of well-orderings of infinite subsets of $X$ into $X$ with the property that if $Y \subseteq X$ is infinite and $R$ well-orders $Y$, then $F(R) \in X \setminus Y$.

Assume for now that we have such a function $F$. By assumption, $\omega \leq X$, so $\text{dom}(F) \neq \emptyset$. Starting with any element of $\text{dom}(F)$, we can then argue as in the proof of Zermelo’s theorem, to produce a well-ordering $W$ of all of $X$. This is of course a contradiction (since then $F(X) \in \emptyset$).

[In more detail: Given $<_Y$ a well-ordering of an infinite subset $Y$ of $X$, we have $F(<_Y) \notin Y$, which provides us with a larger well-ordering, namely, put $F(<_Y)$ on top of $(Y, <_Y)$. Starting with a $(Y, <)$ witnessing $\omega \leq X$, we then obtain a well-ordering $(W, <)$ with $W \subseteq X$ and $W$ largest among the well-orderable subsets of $X$ extending $Y$ such that for all $x \in W \setminus Y$, $F(\{y \in W : y < x\}) = x$. Since $F(<_Z) \notin Z$ for any infinite well-order $(Z, <_Z)$, the map $F$ is injective, and we reach a contradiction as indicated above.]

Now we proceed to find $F$:
Lemma 20 There is a canonical way of associating to a well-ordering $<$ of an infinite set $Y$ a bijection $H : Y \to \text{Seq}(Y)$.

Proof: Begin by noticing:

Lemma 21 There is a canonical bijection $\alpha \to \alpha \times \alpha$ for each infinite ordinal $\alpha$.

Proof: Recall that Gödel’s pairing well-orders $\text{ORD} \times \text{ORD}$ in order-type $\text{ORD}$: Given ordinals $\eta, \beta, \gamma, \delta$, set $(\eta, \beta) <_G (\gamma, \delta)$ iff

- $\max\{\eta, \beta\} < \max\{\gamma, \delta\}$, or
- $\max\{\eta, \beta\} = \max\{\gamma, \delta\}$ and $(\eta, \beta) <_{\text{lex}} (\gamma, \delta)$, where $<_{\text{lex}}$ is the lexicographic ordering.

One easily checks that $<_G$ is a well-ordering of $\text{ORD} \times \text{ORD}$ that is set-like, meaning that the set of $<_G$ predecessors of any pair of ordinals is a set. In fact, something stronger holds. Say that an infinite ordinal $\beta$ is an indecomposable ordinal iff whenever $\gamma, \delta < \beta$, then $\gamma + \delta < \beta$. One then shows:

- $\gamma + \beta = \beta$ for all $\gamma < \beta$; in fact, this is equivalent to the indecomposability of $\beta$.
- There is some ordinal $\gamma < \beta$ such that $\beta = \omega^\gamma$, where exponentiation is in the ordinal sense. (Equality is possible, in which case $\beta$ is called an $\epsilon$ number.) In fact, this is also equivalent to the indecomposability of $\beta$.
- For any $\gamma, \delta < \beta$, $\text{ot}((\gamma, \delta), <_G) < \beta$ and therefore $\text{ot}(\beta \times \beta, <_G) = \beta$.

These statements are all verified by transfinite induction. The last one provides us with an explicit (canonical) bijection between $\alpha$ and $\alpha \times \alpha$ whenever $\alpha$ is indecomposable.

To prove the result in general, notice that for any infinite ordinal $\alpha$ there is a largest ordinal $\beta$ such that $\alpha \geq \omega^\beta$ and, for this $\beta$, there is a largest positive $n \in \omega$ with $\alpha \geq \omega^\beta \cdot n$. There is then a unique $\gamma < \omega^\beta$ such that $\alpha = \omega^\beta \cdot n + \gamma$. There is an obvious bijection between $\omega^\beta + \gamma$ and $\gamma + \omega^\beta = \omega^\beta$. By induction on $n$, from the bijection shown above, there is an explicit bijection between $\omega^\beta$ and $\omega^\beta \cdot n$. It follows that there is an explicit (canonical) bijection between $\alpha$ and $\omega^\beta$ and, therefore, between $\alpha$ and $\alpha \times \alpha$, as we wanted to show. □

By induction, this gives a canonical bijection $\alpha \to \omega^\alpha$ for each $0 < n < \omega$ and, by canonicity, we then have that $\alpha \leq \text{Seq}(\alpha) = \bigcup_n \omega^\alpha$, $\omega \cdot \alpha \sim \omega$, $\omega \cdot \alpha \leq \alpha \cdot \omega \leq \alpha \sim \omega$, and we are done since the Schröder-Bernstein theorem has a constructive proof. □

For $(Y, <)$ an infinite well-order with $Y \subseteq X$ and $H$ as above, let

$$D = \{x \in Y : G^{-1}(H(x)) \text{ is defined and } x \notin G^{-1}(H(x))\}.$$  

Since $D$ is a subset of $X$, $G(D)$ is defined and may or may not belong to $\text{Seq}(Y)$. If it does, then there is an $x_0 \in Y$ such that $G(D) = H(x_0)$, and it follows that $x_0 \in D$ iff
\(x_0 \notin D\), by definition of \(D\). This is a contradiction, and therefore we must conclude that \(G(D) \notin \text{Seq}(Y)\). Hence, \(G(D)\) is a finite sequence and at least one of its elements is not a member of \(Y\).

Define \(F(\langle \rangle)\) to be the element of this sequence with least index which is not in \(Y\). \(\square\)

**Definition 22** A set \(X\) is Dedekind-finite (or D-finite) iff \(X\) is not equipotent with any of its proper subsets. Equivalently, any injection \(f: X \rightarrow X\) is a bijection.

Of course, under choice, a set is Dedekind-finite if and only if it is finite, but without choice one may have infinite Dedekind-finite sets. Notice that if \(\omega \leq X\), as in the hypothesis of the Halbeisen-Shelah result, then \(X\) is Dedekind-infinite, since \(\omega\) is.

In abstract algebra, a ring \(R\) is called Dedekind-finite iff for any \(a, b \in R\), \(ab = 1\) implies \(ba = 1\). This has nothing to do with the set theoretic notion.

**Homework problem 1.** Show Specker’s result that if \(m \geq 5\) then \(2^m \nleq m^2\) and conclude that if \(m\) is infinite then, for all finite \(n\), one has \(n \cdot m < 2^m\).

### 7 Hartogs theorem

For any set \(X\) define \(\aleph(X) = \{\alpha : \exists f: \alpha \rightarrow X \text{ (f is 1-1)}\}\). \(\aleph(\cdot)\) is called Hartogs function.

**Theorem 23** For all sets \(X\), \(\aleph(X)\) is a set. In fact, it is an ordinal. In fact, it is a cardinal (i.e., an initial ordinal).

**Proof:** That \(\aleph(X)\) is a set follows from noticing that it is the surjective image of \(\mathcal{P}(X \times X)\) under the map \(\pi\) that sends \(R\) to zero unless \(R\) is a well-ordering of its field, in which case \((\text{field}(R), R) \cong \alpha\) for a unique ordinal \(\alpha\), and we set \(\pi(R) = \alpha\). That this is a surjection follows from noticing that any injection \(f: \alpha \rightarrow X\) induces an \(R\) with \(\pi(R) = \alpha\).

Now that we know that \(\aleph(X)\) is a set of ordinals, to show that it is an ordinal it suffices to notice that it is closed under initial segments, but this is clear: If \(f: \alpha \rightarrow X\) witnesses \(\alpha \in \aleph(X)\), then for any \(\beta < \alpha\), the map \(f \upharpoonright \beta\) witnesses \(\beta \in \aleph(X)\). Similarly, that \(\aleph(X)\) is a cardinal follows from noticing that if \(\alpha \sim \beta\) and \(\alpha \in \aleph(X)\) then \(\beta \in \aleph(X)\). \(\square\)

We have shown that for any set \(X\) there is a first ordinal (necessarily, a cardinal) that does not inject into \(X\). For example, if there is no \(\omega_1\)-sequence of distinct reals (i.e., if \(\omega_1 \nleq \mathbb{R}\)) as is the case under determinacy, then \(\aleph(\mathbb{R}) = \omega_1\). Of course, under choice, \(\aleph(X) = |X|^+\).

**Theorem 24** \(\aleph(X) \leq \mathcal{P}^3(X)\).

**Proof:** For \(f: \alpha \rightarrow X\) an injection, let \(T_f\) be the set

\[
\{f[\gamma] : \gamma \leq \alpha\} = \{f(\beta) : \beta < \gamma\} : \gamma \leq \alpha\}
\]
and notice that \( T_f \) uniquely determines \( f \) (otherwise, reach a contradiction by considering the first \( \gamma \) such that \( T_f \) does not uniquely determine \( f \mid \gamma \)).

For \( \alpha \in \aleph(X) \) let \( A_\alpha = \{ f \in \alpha : f \text{ is } 1-1 \} \) and set \( \pi(\alpha) = \{ T_f : f \in A_\alpha \} \). This is an injection of \( \aleph(X) \) into \( \mathcal{P}^3(X) \). □

One cannot improve the result to \( \aleph(X) \preceq \mathcal{P}(X) \). For example, \( \aleph(\omega) = \aleph_1 \) does not necessarily embed into \( \mathcal{P}(\omega) \sim \mathbb{R} \). However, \( \omega_1 \preceq \mathcal{P}^2(\omega) \) since \( \omega_1 \preceq \mathcal{P}(\mathbb{R}) \): simply map \( \alpha < \omega_1 \) to the set of reals coding \( \alpha \) in the fashion explained in Section 3 above.

I don’t think one can improve the result to \( \aleph(X) \preceq \mathcal{P}(X) \), but I don’t know of any references for this, and counterexamples are more difficult to come by. To illustrate the difficulty, we prove the following simple lemma; notice that if \( X \sim X^2 \), then \( X \) satisfies the hypothesis, and that if \( X \) is Dedekind-finite, then \( X \not\sim X^2 \).

**Lemma 25** Suppose that either there is \( Y \) such that \( X \sim Y^2 \) or \( X \) is Dedekind-finite. Then \( \aleph(X) \preceq \mathcal{P}^2(X) \).

**Proof:** Assume first that \( X \) is Dedekind-finite. We may as well assume that \( X \) is infinite, or there is nothing to prove. As shown in the next section, \( \aleph(X) = \omega \), and \( \omega \preceq \mathcal{P}^2(X) \).

For any \( Y \), the map \( \alpha \mapsto \{ R \subseteq Y^2 : R \text{ is a well-ordering (of its field) isomorphic to } \alpha \} \) is an injection of \( \aleph(Y) \) into \( \mathcal{P}^2(Y^2) \). Suppose now that \( X \sim Y^2 \). If \( Y \) is finite, the result is obvious; otherwise, it follows from the observation that \( \aleph(Y) = \aleph(Y^2) \) holds for any infinite set \( Y \) (this is homework problem 3, below). □

Notice that under choice the assumption of the lemma always holds, but in this case we in fact have \( \aleph(X) \preceq \mathcal{P}(X) \).

The axiom of choice is equivalent to the statement: For all ordinals \( \alpha \), \( \mathcal{P}(\alpha) \) is well-orderable; a proof can be found below. Hence, if choice fails, some \( \mathcal{P}(\alpha) \) is not well-orderable, and by the above, \( \aleph(\alpha) = |\alpha|^+ \preceq \mathcal{P}^2(\alpha) \). In Section 10 we will show that if in addition the continuum hypothesis holds for \( \alpha \), then \( \aleph(\mathcal{P}(\alpha)) = |\alpha|^+ \) as well. An example of this situation occurs under determinacy, where every set of reals has the perfect set property, so the continuum hypothesis holds.

Here are some corollaries of the injection \( \aleph(X) \preceq \mathcal{P}^3(X) \):

**Corollary 26** For any set \( X \), \( \aleph(X) < |\mathcal{P}^4(X)| \) and \( \aleph(X) < \aleph(\mathcal{P}^3(X)) \). □

**Corollary 27** There is no sequence \( (X_n : n < \omega) \) such that for all \( n \), \( |\mathcal{P}(X_{n+1})| \leq |X_n| \).

**Proof:** Otherwise, by the previous corollary, \( (\aleph(X_{3n}) : n < \omega) \) would be a strictly decreasing sequence of ordinals. □

Remember that the sequence \( (\aleph_\alpha : \alpha \in \text{ORD}) \) of infinite initial ordinals can be defined in terms of Hartogs function:

- \( \aleph_0 = \omega_0 = \omega \).
• Given $\aleph_\alpha = \omega_\alpha$, let $\aleph_{\alpha+1} = \omega_{\alpha+1} = \aleph(\aleph_\alpha)$.

• At limit stages, simply set $\aleph_\alpha = \omega_\alpha = \sup_{\beta < \alpha} \omega_\beta$, where the supremum coincides with the union of the corresponding ordinals.

Above, we showed that $\aleph(\aleph_\alpha) \preceq P^2(\aleph_\alpha)$ if $\aleph_\alpha \sim \aleph_\alpha^2$ for some $\aleph_\alpha$, or if $\aleph_\alpha$ is Dedekind-finite.

**Theorem 28** The axiom of choice is equivalent to the statement that any Dedekind-infinite cardinal is a square.

**Proof:** Let $X$ be a set. Assuming that every D-infinite cardinal is a square, we need to show that $X$ is well-orderable. We may assume that $\omega \preceq X$. Otherwise, replace $X$ with $X \cup \omega$.

Let $\kappa = \aleph(X)$. Assume that $X \cup \kappa$ is a square, say $X \cup \kappa \sim \aleph_\alpha^2$. Then $\kappa \preceq \aleph_\alpha^2$. By Homework problem 2, $\kappa \preceq \aleph_\alpha$, so $\aleph_\alpha \sim \aleph_\alpha \cup \aleph_\alpha$ for some $\aleph_\alpha$, and $X \cup \kappa \sim \aleph_\alpha^2 \sim \kappa^2 \cup 2 \times \kappa \times \aleph_\alpha \preceq \kappa \times \aleph_\alpha$.

**Lemma 29** Suppose $A, B, C$ are D-infinite sets and $\lambda$ is an (infinite) initial ordinal. If $\lambda \times A \preceq B \cup C$ then either $\lambda \preceq B$ or $A \preceq C$.

**Proof:** Let $f : \lambda \times A \to B \cup C$ be an injection. If there is some $a \in A$ such that $f(\cdot, a) : \lambda \to B$ we are done, so we may assume that for all $a \in A$ there is some $\alpha \in \lambda$ such that $f(\alpha, a) \in C$. Letting $\alpha_a$ be the least such $\alpha$, the map $a \mapsto f(\alpha_a, a)$ is an injection of $A$ into $C$. □

By the lemma, it must be that either $\kappa \preceq X$ or else $Z \preceq \kappa$. The former is impossible since $\kappa = \aleph(X)$, so $Z$ is well-orderable, and thus so is $Y$, and since $Y \sim Y^2 \succeq X$, then $X$ is well-orderable as well. □

8 Dedekind-finite sets

**Theorem 30** A set $X$ is D-finite iff $\aleph(X) \preceq \omega$.

**Proof:** Clearly, $\omega \preceq X$ implies that $X$ is Dedekind-infinite. If $f : X \to X$ is injective but not surjective, and $a \in X \setminus f''X$, then the sequence $(f^n(a) : n \in \omega)$ witnesses that $\omega \preceq X$. □

**Corollary 31** Let $X$ be a set and let $0 < n < \omega$. Then $X$ is D-finite iff $n \cdot X$ is D-finite.

**Proof:** Clearly, every subset of a D-finite set is D-finite. To see the other direction, if $\omega$ injects into $\bigcup_{i < n} X$ then it injects into one of the copies of $X$ (by the pigeonhole principle). □

**Corollary 32** If $X$ is infinite, then $P^2(X)$ is Dedekind-infinite.
Proof: $\omega \leq \mathcal{P}^2(X)$ as witnessed by $n \mapsto A_n$, where $A_n = [X]^n = \{Y \subseteq X : |Y| = n\}$. The assumption that $X$ is infinite is used to check (by induction) that no $A_n$ is empty, which immediately gives us that the map is injective. □

Homework problem 2. Show that if $\kappa$ is a well-ordered cardinal (an initial ordinal) and $\kappa \leq A \times B$, then either $\kappa \leq A$ or $\kappa \leq B$. In particular, if $A$ and $B$ are Dedekind-finite, then so is $A \times B$.

Homework problem 3. Show that for any infinite set $A$, $\aleph(A) = \aleph(A^2)$.

9 $k$-trichotomy

This is a very recent result (from 2008) by Feldman and Orhon, see [3], available at the arXiv.

Definition 33 Trichotomy is the statement that any two sets are comparable.

Theorem 34 (Hartogs) Trichotomy is equivalent to choice.

Proof: If choice holds, any set is equipotent with an ordinal, and any two ordinals are comparable.

Conversely, if any two sets are comparable, in particular $X$ and $\aleph(X)$ are. Since $\aleph(X) \nleq X$, we must have $X \leq \aleph(X)$, and we are done. □

Definition 35 Let $2 \leq k < \omega$. $k$-trichotomy is the statement that given any $k$ sets, at least two are comparable.

Theorem 36 (Feldman-Orhon) For any $2 \leq k < \omega$, $k$-trichotomy is equivalent to choice.

The argument below is due to Blass.

Proof: Assume $k$-trichotomy. Let $X$ be a set. We want to show that $X$ is well-orderable. Let $Q(X) = X \times \mathcal{P}(X)$.

Lemma 37 Let $\kappa$ be an initial ordinal. If $Q(X) \leq X + \kappa$, then $X$ is well-orderable.

Here, $+$ is to be understood as disjoint union.

Proof: Let $f : Q(X) \rightarrow X + \kappa$ be 1-1. For each $x$, since $\mathcal{P}(X)$ does not inject into $X$, it must be the case that the range of $f(x, \cdot)$ meets $\kappa$. Let $\alpha_x$ be the least member of this intersection. Then $x \mapsto \alpha_x$ is an injection of $X$ into $\kappa$. □
Note that if \( X \sim X + X \), then \( \mathcal{P}(X) \leq Q(X) \leq \mathcal{P}(X) \times \mathcal{P}(X) \sim \mathcal{P}(X + X) \sim \mathcal{P}(X) \), and it follows that \( \mathcal{P}(X) \sim Q(X) \). However, it is not necessarily the case that \( X \sim X + X \) for \( X \) infinite. For example, this fails if \( X \) is D-finite.

**Question.** If \( \mathcal{P}(X) \leq X + \kappa \) for some initial ordinal \( \kappa \), does it follow that \( X \) is well-orderable?

Define a sequence of \( k \) cardinals, \( \kappa_0(X,k), \ldots, \kappa_{k-1}(X,k) \) as follows:

- \( \kappa_0 = \kappa_0(X,k) = \aleph(Q^{k-1}(X)) \).
- \( \kappa_{i+1} = \kappa_{i+1}(X,k) = \aleph(Q^{k-i-1}(X) + \kappa_i) \).

(Then \( \kappa_0 < \kappa_1 < \ldots \))

We claim that if \( X \) is not well-orderable, the set \( \{ Q^{k-i-1}(X) + \kappa_i : i < k \} \) contradicts \( k \)-trichotomy. Or, positively, \( k \)-trichotomy applied to this set implies the well-orderability of \( X \). Because if \( i \neq j \) and there is an injection \( Q^{k-i-1}(X) + \kappa_i \rightarrow Q^{k-j-1}(X) + \kappa_j \), then in particular \( \kappa_i \leq Q^{k-j-1}(X) + \kappa_j \). If \( j = k - 1 \), then \( i < j \). Otherwise, \( \kappa_{j+1} \) is defined, and \( \kappa_i < \kappa_{j+1} \) (since \( \kappa_{j+1} \) is the aleph of the set that \( \kappa_i \) injects into). Since the sequence of cardinals \( \kappa_n \) is increasing, we must have \( i < j + 1 \), so \( i \leq j \) and (since \( i \neq j \)) in fact \( i < j \).

Let \( Y = Q^{k-j-1}(X) \). Then \( X \preceq Y < Q(Y) \preceq Q^{j-i}(Y) \), using that \( j > i \). Also, \( Q^{j-i}(Y) = Q^{k-i-1}(X) \preceq Q^{k-i-1}(X) + \kappa_i \preceq Q^{k-j-1}(X) + \kappa_j \), by assumption. But \( Q^{k-j-1}(X) + \kappa_j = Y + \kappa_j \), and we have that \( Q(Y) \preceq Y + \kappa_j \). By the lemma, \( Y \) is well-orderable and we are done, since \( X \) injects into \( Y \). \( \square \)

**Open question.** With the obvious definition, does \( \omega \)-trichotomy imply choice?

Let \( \infty \)-trichotomy be the statement that any infinite family of sets contains two that are comparable.

**Open question.** Does \( \omega \)-trichotomy imply \( \infty \)-trichotomy? Does \( \infty \)-trichotomy imply choice?

## 10 The generalized continuum hypothesis

The last topic of this chapter is Specker’s theorem that GCH implies AC.

**Definition 38** Let \( m \) be a cardinal. The continuum hypothesis for \( m \), \( \text{CH}(m) \), is the statement that for any cardinal \( n \), if \( m \leq n \leq 2^m \), then either \( n = m \) or else \( n = 2^m \).

**Definition 39** The Generalized Continuum Hypothesis, \( \text{GCH} \), is the statement that \( \text{CH}(m) \) holds for all infinite cardinals \( m \).

**Theorem 40** (Specker) \( \text{GCH} \) implies \( \text{AC} \), the axiom of choice.
In fact, it suffices that GCH holds for all infinite initial ordinals. This was shown independently by Herman Rubin and Arthur Kruse in 1960; we will obtain it as a consequence of a result of Kanamori and Pincus in [7].

Specker’s result is in fact a local argument. What one proves is the following:

**Theorem 41 (Specker)** Assume CH(m) and CH(2^m). Then 2^m is well-orderable and, in fact, 2^m = ℵ(m).

It follows immediately from the theorem that GCH implies AC, since the result gives that any (infinite) m embeds into ℵ(m).

**Proof:** The argument depends on 2 lemmas.

**Lemma 42** CH(m) implies that m + m = m^2 = m.

**Proof:**

a) m ≤ m + 1 < 2^m, where the strict inequality is Specker’s Lemma 18. By CH(m), we have m = m + 1. (In particular, m is Dedekind-infinite.)

b) m ≤ m + m ≤ 2^m + 2^m = 2^{m+1} = 2^m, by a). By CH(m), we have m + m = m or m + m = 2^m.

c) If m + m = 2^m, we can inject \( P(X) \) into the set Seq(\( X \)) of finite sequences of elements of \( X \). For example, if \( P(X) \) is the disjoint union of A and B, both of size m, there are bijections \( \pi : A \to X \) and \( \rho : B \to X \), and we can use them to define an injection \( j : P(X) \to Seq(X) \) as follows: Fix \( a \in X \). Define \( j \upharpoonright A : A \to X^1 \) by \( j(r) = (\pi(r)) \), and \( j \upharpoonright B : B \to X^2 \) by \( j(r) = (a, \rho(r)) \).

d) But there is no such injection, by the Halbeisen-Shelah Theorem 19. Recall that the theorem has the hypothesis that \( X \) is Dedekind-infinite, that we showed in a).

e) So m + m = m.

f) Finally, m ≤ m × m ≤ 2^m × 2^m = 2^{m+m} = 2^m, by e). By CH(m), it follows that m × m = m, or m × m = 2^m.

g) But m × m = 2^m immediately contradicts the Halbeisen-Shelah theorem.

h) So m × m = m. □

**Lemma 43** CH(m) implies that either 2^m is well-orderable (and equal to ℵ(m)), or else ℵ(m) does not inject into 2^m.

**Proof:** Assume CH(m). Suppose ℵ(m) injects into 2^m. Note that m < m + ℵ(m), since ℵ(m) does not inject into m. Thus, m < m + ℵ(m) ≤ 2^m + 2^m = 2^{m+1} = 2^m since CH(m) implies m + 1 = m, as shown in Lemma 42. By CH(m) again, m + ℵ(m) = 2^m.

By CH(m), m + m = m so, if X has size m, then Q(X) = X × P(X) is equipotent with P(X), as shown right after the proof of Lemma 37.

As shown in Lemma 37, if Q(X) injects into X + \( \kappa \) for some ordinal \( \kappa \), then X is well-orderable. It follows that in our situation X is well-orderable, so m is an ordinal. But then m + ℵ(m) is an ordinal and in fact, it equals ℵ(m). Since it also equals 2^m, we are done. □
Using these lemmas, we prove Specker’s result as follows:

- Assuming CH(n) but \(2^n\) not well-orderable, from Lemma 43 it follows that \(\aleph(n) = \aleph(2^n)\).
- Hence if CH(m), CH(2^m), and \(2^m\) is not well-orderable, then neither is \(2^{2^m}\), so \(\aleph(2^m) = \aleph(2^{2^m})\) and \(\aleph(m) = \aleph(2^m)\).
- Hence if \(X\) has size \(m\), then \(\aleph(X) = \aleph(\mathcal{P}(\mathcal{P}(X)))\). But \(X\) and \(X^2\) have the same size by Lemma 42, so \(\aleph(X) = \aleph(\mathcal{P}(\mathcal{P}(X^2)))\).
- But this contradicts that \(\aleph(X)\) injects into \(\mathcal{P}(\mathcal{P}(X^2))\), as shown in Lemma 25.

This completes the proof of Specker’s theorem. □

Open question. (ZF) Does CH(m) imply that \(m\) is well-orderable?

CH(m) does not imply that \(2^m\) is well-orderable. For example, under determinacy, every set of reals has the perfect set property, so CH holds. But \(\mathbb{R}\) is not well-orderable.

Homework problem 4. Suppose that \(m + m = m\) and \(m + n = 2^m\). Then \(n = 2^m\).

Some hypothesis is necessary for the result of Homework problem 4 to hold. For example, it is consistent that there is an infinite Dedekind finite set \(X\) with \(\mathcal{P}(X)\) also Dedekind-finite. (On the other hand, \(\mathcal{P}^2(X)\) is necessarily Dedekind-infinite.) However, if \(2^m\) is Dedekind-finite and \(p + n = 2^m\) for nonzero cardinals \(p, n\), then obviously \(n < 2^m\).

I am not sure who first proved the following; it appears in Kanamori-Pincus [7].

**Theorem 44** Assume CH(m), CH(n) and \(m < n\). Then \(2^m \leq n\). In particular, there is at most one initial ordinal \(\kappa\) such that CH(\(\kappa\)) but \(2^\kappa\) is not well-orderable.

**Proof:** If \(m < n\) then \(n \leq 2^m + n \leq 2^n + n = 2^{n+1} = 2^n\) since \(n + 1 = n\) by CH(n).

Again by CH(n), either \(2^m + n = n\) or \(2^m + n = 2^n\).

If \(2^m + n = 2^n\), then Homework problem 4 implies that \(2^m = 2^n\), so \(m < n < 2^m\) and CH(m) fails.

Hence, \(2^m + n = n\), so \(2^m \leq n\).

The ‘in particular’ clause follows immediately. □

**Corollary 45** GCH for initial ordinals implies AC.

**Proof:** As shown below, AC is equivalent to \(\mathcal{P}(\alpha)\) being well-orderable for all ordinals \(\alpha\).

This follows immediately from GCH for initial ordinals, by Theorem 44. □
**Theorem 46** The axiom of choice is equivalent to the statement that \( P(\alpha) \) is well-orderable for all ordinals \( \alpha \).

**Proof:** Assume in ZF that the power set of any ordinal is well-orderable. We want to conclude that choice holds, i.e., that every set is well-orderable. A natural strategy is to proceed inductively, showing that each \( V_\alpha \) is well-orderable: Clearly, if the result is true, each \( V_\alpha \) would be well-orderable. But also, given any set \( x \), it belongs to some \( V_\alpha \) and, since the latter is transitive, in fact \( x \subseteq V_\alpha \) and therefore \( x \) is well-orderable as well. The strategy is suggested by the fact that for all \( \alpha \), \( V_{\alpha+1} = P(V_\alpha) \), so a well-ordering of \( V_\alpha \) gives us a well-ordering of \( V_{\alpha+1} \) thanks to our initial assumption.

We argue by induction: Clearly \( V_0 \) is well-ordered by the well-ordering \( <_0 = \emptyset \). Given a well-ordering \( < \) of \( V_\alpha \), there is a unique ordinal \( \beta \) and a unique order isomorphism \( \pi : (V_\alpha, <) \to (\beta, \in) \).

By assumption, \( P(\beta) \) is well-orderable, and any well-ordering of it induces (via \( \pi^{-1} \)) a well-ordering of \( V_{\alpha+1} \).

We are left with the task of showing that \( V_\alpha \) is well-orderable for \( \alpha \) limit. The natural approach is to patch together the well-orderings of the previous \( V_\beta \) into a well-ordering of \( V_\alpha \). This approach meets two obstacles.

The first one is that the well-orderings of different \( V_\beta \) are not necessarily compatible, so we need to be careful on how we “patch them together.” This is not too much of an obstacle: The natural solution is to simply order the sets as they appear inside \( V_\alpha \). More precisely, define \( x < y \) for \( x, y \in V_\alpha \), iff

- Either \( \text{rk}(x) < \text{rk}(y) \), or else
- \( \text{rk}(x) = \text{rk}(y) = \beta \), say, and if \( <_{\beta+1} \) is the well-ordering of \( V_{\beta+1} \), then \( x <_{\beta+1} y \).

It is easy to see that this is indeed a well-ordering of \( V_\alpha \): Given a non-empty \( A \subseteq V_\alpha \), let \( \gamma \) be least so that \( A \) has an element of rank \( \gamma \). Then the \( <_{\gamma+1} \)-first among these elements would be the \( < \)-least element of \( A \).

The second obstacle is more serious. Namely, the assumption is simply that there is a well-ordering of each \( P(\delta) \), not that there is any canonical way of choosing one. In order for the argument above to work, we need not just that each \( V_\beta \) for \( \beta < \alpha \) is well-orderable, but in fact we need to have selected a sequence \( (<_{\beta+1} : \beta < \alpha) \) of well-orderings of the \( V_{\beta+1} \), with respect to which we proceeded to define the well-ordering \( < \) of \( V_\alpha \).

The way we began the proof suggests a solution: When we argued that it suffices to well-order each \( V_\gamma \), we considered an arbitrary set \( x \) and noticed that if \( x \subseteq V_\beta \), then a well-ordering of \( V_\beta \) gives us a well-ordering of \( x \). Similarly, given \( \alpha \) limit, if we can find \( \delta \) large enough so each \( |V_\beta| \) for \( \beta < \alpha \) is below \( \delta \), then we can use a well-ordering of \( P(\delta) \) to induce the required well-ordering \( <_\beta \).

We now proceed to implement this idea: Let \( \delta = \sup_{\beta < \alpha} |V_\beta|^+ \). (Notice that this makes sense since, inductively, each \( V_\beta \) with \( \beta < \alpha \) is well-orderable and therefore isomorphic to
Let $<^*$ be a well-ordering of $\mathcal{P}(\delta)$. We use $<^*$ to define a sequence $(<^*_\beta: \beta < \alpha)$ so that $<^*_\beta$ well-orders $V_\beta$ for all $\beta < \alpha$. We use recursion on $\beta < \alpha$ to define this sequence. Again, $<_0 = \emptyset$. At limit stages $\gamma < \alpha$ we copy the strategy with which we tried to well-order $V_\alpha$ to define $<^*_\gamma$: For $x, y \in V_\gamma$, set $x <^*_\gamma y$ iff

- Either $\text{rk}(x) < \text{rk}(y)$, or else
- $\text{rk}(x) = \text{rk}(y) = \beta$, say, and $x <^*_\beta y$.

Finally, given $<^*_\beta$, we describe how to define $<^*_\beta + 1$: Let $\xi = \xi_\beta$ be the unique ordinal such that there is an order isomorphism $\pi: (V_\beta, <^*_\beta) \to (\xi, \in)$. Since $|\xi| = |V_\beta|$, then $\xi < \delta$, so $\xi \subset \delta$ and the well-ordering $<^*$ of $\mathcal{P}(\delta)$ also well-orders $\mathcal{P}(\xi)$. Via $\pi^{-1}$, this induces the well-ordering $<^*_\beta + 1$ of $V_{\beta + 1}$ we were looking for. □

By Specker’s result, if CH($m$) but $2^m$ is not well-orderable, then CH($2^m$) must fail. Kanamori and Pincus proved that in this case in fact there is an increasing sequence of cardinals of length $\text{cf}(\aleph(m))$ between $2^m$ and $2^{2^m}$.

Here is a rough idea of their argument: Recall that $\mathcal{O}(X)$ denotes the set of well-orderings of subsets of $X$, so $\mathcal{O}(X) \subseteq \mathcal{P}(X \times X)$. Define $\mathcal{O}^\xi(X)$ for ordinals $\xi$ by iterating $\mathcal{O}$ transfinitely, taking unions at limit stages and setting $\mathcal{O}^0(X) = X$. By induction, one shows the following three facts:

- For any set $X$ and ordinal $\xi$, $\xi \leq \mathcal{O}^\xi(X)$.
- For any infinite $X$ and any $\xi$, $\mathcal{O}^\xi(X) \leq \mathcal{P}(X \times \aleph(\mathcal{O}^\xi(X)))$.
- For any ordinals $\alpha < \beta$ and any set $X$, $|\mathcal{O}^\alpha(X)| < |\mathcal{O}^\beta(X)|$.

Also, one easily verifies:

- $\aleph(X) \leq \aleph(Y)$ implies $\aleph(\mathcal{O}(X)) \leq \aleph(\mathcal{O}(Y))$. In particular, if $\aleph(X) = \aleph(Y)$, then $\aleph(\mathcal{O}(X)) = \aleph(\mathcal{O}(Y))$.

Assume now that CH($m$) holds but $2^m$ is not well-orderable. Write $\mathcal{O}^\xi(m)$ for the cardinality of $\mathcal{O}^\xi(X)$ for any set $X$ of cardinality $m$. Notice:

- $2^m = \aleph(m)$ and so $\aleph(m) = \aleph(\mathcal{O}(m))$.

Since $\aleph(\mathcal{O}^\xi(X)) \geq |\xi|^+$, there must be a least ordinal $\gamma$ such that $\aleph(m) < \aleph(\mathcal{O}^\gamma(m))$. This $\gamma$ is necessarily an infinite limit ordinal. One concludes the proof by showing:

- For any $\xi$ with $1 < \xi < \gamma$, one has $2^m < \mathcal{O}^\xi(m) \leq 2^{2^m}$.
- $\gamma \geq \text{cf}(\aleph(m))$.  

20
A. Appendix.

I would like to review the results relating $\aleph(X)$ and $\mathcal{P}^2(X)$ in a way that gives some additional information.

**Definition 47** For an infinite set $X$, set $\aleph^*(X) = \{ \alpha : \exists f : X \rightarrow \alpha \text{ surjective} \}$ and $\aleph_*(X) = \aleph^*(X) \cup \{ 0 \}$.

**Lemma 48** $\aleph^*(X)$ is a set and $\aleph_*(X)$ is an initial ordinal.

**Proof:** Suppose that $\pi : X \rightarrow Y$ is a surjection. Define $E_\pi \subseteq X^2$ by $x E_\pi y$ iff $\pi(x) = \pi(y)$. Clearly, $E_\pi$ is an equivalence relation on $X$, and there is an obvious bijection $[x]_{E_\pi} \mapsto \pi(x)$ between the quotient $X/E_\pi$ and $Y$. The collection of quotients is in bijection with the collection of equivalence relations, that is clearly a set. From the above it follows that an ordinal $\alpha$ is the surjective image of $X$ iff there is an equivalence relation on $X$ whose quotient is well-orderable and has size $|\alpha|$. This shows that $\aleph^*(X)$ is a set and that if $\aleph_*(X)$ is an ordinal, then it is an initial ordinal. But this follows from noticing that if $0 < \beta < \alpha$ and $\pi : X \rightarrow \alpha$ is surjective, then $\pi^* : X \rightarrow \beta$ is surjective, where $\pi^*(x) = \pi(x)$ if $\pi(x) < \beta$, and $\pi^*(x) = 0$ otherwise. □

**Lemma 49** $\aleph(X) \leq \aleph_*(X)$.

**Proof:** If $g : \alpha \rightarrow X$ is injective and $0 < \alpha$, let $f : X \rightarrow \alpha$ be the function $f(x) = \beta$ if $x = g(\beta)$ and $f(x) = 0$ otherwise. Then $f$ is surjective. □

It is possible that $\aleph(X) = \aleph_*(X)$. For example, this holds if $X$ is well-orderable. It is also possible that $\aleph(X) < \aleph_*(X)$. For example, under determinacy $\aleph(\mathbb{R}) = \omega_1$ but $\Theta = \aleph_*(\mathbb{R})$ is much larger.

**Lemma 50** If $\pi : X \rightarrow Y$ is surjective, then $\pi^{-1} = \mathcal{P}(Y) \rightarrow \mathcal{P}(X)$ is injective. □

**Corollary 51** $\mathcal{P}(\aleph(X)) \leq \mathcal{P}^2(X \times X)$.

**Proof:** Define a map $\pi : \mathcal{P}(X \times X) \rightarrow \aleph(X)$ by setting $\pi(R) = 0$ for a relation $R \subseteq X \times X$ unless $R$ is a well-ordering of its field in order type $\alpha$, in which case $\pi(R) = \alpha$. This is a surjection. □

**Corollary 52** $\mathcal{P}(\alpha^+) \leq \mathcal{P}^2(\alpha)$ for all ordinals $\alpha$.

**Proof:** This is clear by induction if $\alpha$ is finite. Otherwise, $\alpha \sim \alpha \times \alpha$. □

**Corollary 53** $\mathcal{P}(\aleph(X)) \leq \mathcal{P}^2(X)$ if $X \sim Y^2$ for some $Y$. 

21
Proof: Again, it suffices to consider infinite $X$. In this case, $\aleph(X) = \aleph(Y)$ by Homework problem 3. □

Corollary 54 $\aleph(\mathcal{P}^2(X)) \geq \sup\{\aleph(\mathcal{P}^2(\alpha)) : \alpha \in \aleph_*(X)\} \geq \sup\{\aleph(\mathcal{P}(\alpha^+)) : \alpha \in \aleph_*(X)\}$.

Proof: If $\alpha < \aleph_*(X)$ then there is a surjection from $\mathcal{P}(X)$ onto $\mathcal{P}(\alpha)$. This is clear if $\alpha = 0$. Otherwise, if $\pi : X \to \alpha$ is onto, consider $\hat{\pi} : \mathcal{P}(X) \to \mathcal{P}(\alpha)$ given by $\hat{\pi}(A) = \pi[A]$. Hence, there is an injection of $\mathcal{P}^2(\alpha)$ into $\mathcal{P}^2(X)$. The result follows. □

Corollary 55 $\mathcal{P}(\aleph(X)) \leq \mathcal{P}^2(X)$ if $\aleph(X)$ is a successor cardinal. □

Corollary 56 Suppose that $\kappa > 0$ is the $\kappa$-th initial ordinal. If $\kappa = \aleph(X)$, then $\mathcal{P}(\aleph(X)) \leq \mathcal{P}^2(X)$.

Proof: Let $\tau$ be a bijection between the set of initial ordinals below $\kappa$ and $\kappa$. The map $\pi : \mathcal{P}(X) \to \kappa$ given by $\pi(A) = \tau(|A|)$ if $A$ is well-orderable, and $\pi(A) = 0$ otherwise, is a surjection. □

Corollary 57 If $X$ is infinite and Dedekind-finite, then $\mathcal{P}(\omega) \leq \mathcal{P}^2(X)$.

Proof: $\aleph(X) = \omega$ is the $\omega$-th initial ordinal. □

Corollary 58 Suppose that $\aleph(X) = \aleph(\mathcal{P}^2(X))$. Then $X$ is not equipotent to a square and $\aleph(X) = \aleph_*(X) = \aleph_\lambda$ for some (infinite) limit ordinal $\lambda < \aleph_\lambda$. □

Notice that, under choice, $\aleph(X)$ is always a successor cardinal; I do not know if the conclusion of the corollary is consistent.

References


*Typeset using LaTeX2WP.*