Appendix B. Zorn's Lemma and the Hausdorff Maximal Principle

Definition B.1. A partial order $\leq$ on X is a relation with following properties

(i) If $x \leq y$ and $y \leq z$ then $x \leq z$.
(ii) If $x \leq y$ and $y \leq x$ then $x = y$.
(iii) $x \leq x$ for all $x \in X$.

Example B.2. Let Y be a set and $X = \mathcal{P}(Y)$. There are two natural partial orders on X.

(1) Ordered by inclusion, $A \leq B$ is $A \subset B$ and
(2) Ordered by reverse inclusion, $A \leq B$ if $B \subset A$.

Definition B.3. Let $(X, \leq)$ be a partially ordered set we say X is **linearly** a **totally** ordered if for all $x, y \in X$ either $x \leq y$ or $y \leq x$. The real numbers $\mathbb{R}$ with the usual order $\leq$ is a typical example.

Definition B.4. Let $(X, \leq)$ be a partial ordered set. We say $x \in X$ is a **maximal** element if for all $y \in X$ such that $y \geq x$ implies $y = x$, i.e. there is no element larger than x. An **upper bound** for a subset E of X is an element $x \in X$ such that $x \geq y$ for all $y \in E$.

Example B.5. Let

$$X = \left\{ \ a = \{1\} \quad b = \{1,2\} \quad c = \{3\} \quad d = \{2,4\} \quad e = \{2\} \ \right\}$$

ordered by set inclusion. Then b and d are maximal elements despite that fact that $b \nleq a$ and $a \nleq b$. We also have

- If $E = \{a, e, c\}$, then E has **no** upper bound.

Definition B.6.

- If $E = \{a, e\}$, then b is an upper bound.
- $E = \{e\}$, then b and d are upper bounds.

Theorem B.7. *The following are equivalent.*

(1) ***The axiom of choice:*** *to each collection,* $\{X_\alpha\}_{\alpha \in A}$, *of non-empty sets there exists a "choice function,"* $x : A \to \coprod_{\alpha \in A} X_\alpha$ *such that* $x(\alpha) \in X_\alpha$ *for all* $\alpha \in A$, *i.e.* $\prod_{\alpha \in A} X_\alpha \neq \emptyset$.
(2) ***The Hausdorff Maximal Principle:*** *Every partially ordered set has a **maximal** (relative to the inclusion order) linearly ordered subset.*
(3) ***Zorn's Lemma:*** *If X is partially ordered set such that every linearly ordered subset of X has an upper bound, then X has a maximal element.*[51]

Proof. $(2 \Rightarrow 3)$ Let X be a partially ordered subset as in 3 and let $\mathcal{F} = \{E \subset X : E \text{ is linearly ordered}\}$ which we equip with the inclusion partial ordering. By 2. there exist a maximal element $E \in \mathcal{F}$. By assumption, the linearly ordered set E has an upper bound $x \in X$. The element x is maximal, for if $y \in Y$ and $y \geq x$,

[51] If X is a countable set we may prove Zorn's Lemma by induction. Let $\{x_n\}_{n=1}^\infty$ be an enumeration of X, and define $E_n \subset X$ inductively as follows. For $n = 1$ let $E_1 = \{x_1\}$, and if E_n have been chosen, let $E_{n+1} = E_n \cup \{x_{n+1}\}$ if x_{n+1} is an upper bound for E_n otherwise let $E_{n+1} = E_n$. The set $E = \cup_{n=1}^\infty E_n$ is a linearly ordered (you check) subset of X and hence by assumption E has an upper bound, $x \in X$. I claim that his element is maximal, for if there exists $y = x_m \in X$ such that $y \geq x$, then x_m would be an upper bound for E_{m-1} and therefore $y = x_m \in E_m \subset E$. That is to say if $y \geq x$, then $y \in E$ and hence $y \leq x$, so $y = x$. (Hence we may view Zorn's lemma as a " jazzed" up version of induction.)

then $E \cup \{y\}$ is still an linearly ordered set containing E. So by maximality of E, $E = E \cup \{y\}$, i.e. $y \in E$ and therefore $y \leq x$ showing which combined with $y \geq x$ implies that $y = x$.[52]

$(3 \Rightarrow 1)$ Let $\{X_\alpha\}_{\alpha \in A}$ be a collection of non-empty sets, we must show $\prod_{\alpha \in A} X_\alpha$ is not empty. Let $\mathcal{G}$ denote the collection of functions $g : D(g) \to \coprod_{\alpha \in A} X_\alpha$ such that $D(g)$ is a subset of A, and for all $\alpha \in D(g)$, $g(\alpha) \in X_\alpha$. Notice that $\mathcal{G}$ is not empty, for we may let $\alpha_0 \in A$ and $x_0 \in X_\alpha$ and then set $D(g) = \{\alpha_0\}$ and $g(\alpha_0) = x_0$ to construct an element of $\mathcal{G}$. We now put a partial order on $\mathcal{G}$ as follows. We say that $f \leq g$ for $f, g \in \mathcal{G}$ provided that $D(f) \subset D(g)$ and $f = g|_{D(f)}$. If $\Phi \subset \mathcal{G}$ is a linearly ordered set, let $D(h) = \cup_{g \in \Phi} D(g)$ and for $\alpha \in D(g)$ let $h(\alpha) = g(\alpha)$. Then $h \in \mathcal{G}$ is an upper bound for Φ. So by Zorn's Lemma there exists a maximal element $h \in \mathcal{G}$. To finish the proof we need only show that $D(h) = A$. If this were not the case, then let $\alpha_0 \in A \setminus D(h)$ and $x_0 \in X_{\alpha_0}$. We may now define $D(\tilde{h}) = D(h) \cup \{\alpha_0\}$ and

$$\tilde{h}(\alpha) = \begin{cases} h(\alpha) & \text{if} \quad \alpha \in D(h) \\ x_0 & \text{if} \quad \alpha = \alpha_0. \end{cases}$$

Then $h \leq \tilde{h}$ while $h \neq \tilde{h}$ violating the fact that h was a maximal element.

$(1 \Rightarrow 2)$ Let $(X, \leq)$ be a partially ordered set. Let $\mathcal{F}$ be the collection of linearly ordered subsets of X which we order by set inclusion. Given $x_0 \in X$, $\{x_0\} \in \mathcal{F}$ is linearly ordered set so that $\mathcal{F} \neq \emptyset$.

Fix an element $P_0 \in \mathcal{F}$. If P_0 is not maximal there exists $P_1 \in \mathcal{F}$ such that $P_0 \subsetneq P_1$. In particular we may choose $x \notin P_0$ such that $P_0 \cup \{x\} \in \mathcal{F}$. The idea now is to keep repeating this process of adding points $x \in X$ until we construct a maximal element P of $\mathcal{F}$. We now have to take care of some details.

We may assume with out loss of generality that $\tilde{\mathcal{F}} = \{P \in \mathcal{F} : P \text{ is not maximal}\}$ is a non-empty set. For $P \in \tilde{\mathcal{F}}$, let $P^* = \{x \in X : P \cup \{x\} \in \mathcal{F}\}$. As the above argument shows, $P^* \neq \emptyset$ for all $P \in \tilde{\mathcal{F}}$. Using the axiom of choice, there exists $f \in \prod_{P \in \tilde{\mathcal{F}}} P^*$. We now define $g : \mathcal{F} \to \mathcal{F}$ by

$$g(P) = \begin{cases} P & \text{if} \quad P \text{ is maximal} \\ P \cup \{f(x)\} & \text{if} \quad P \text{ is not maximal.} \end{cases} \tag{B.1}$$

The proof is completed by Lemma B.8 below which shows that g must have a fixed point $P \in \mathcal{F}$. This fixed point is maximal by construction of g. ■

Lemma B.8. *The function $g : \mathcal{F} \to \mathcal{F}$ defined in Eq. (B.1) has a fixed point.*[53]

Proof. The **idea of the proof** is as follows. Let $P_0 \in \mathcal{F}$ be chosen arbitrarily. Notice that $\Phi = \left\{g^{(n)}(P_0)\right\}_{n=0}^{\infty} \subset \mathcal{F}$ is a linearly ordered set and it is therefore easily verified that $P_1 = \bigcup_{n=0}^{\infty} g^{(n)}(P_0) \in \mathcal{F}$. Similarly we may repeat the process to

[52] Similalry one may show that $3 \Rightarrow 2$. Let $\mathcal{F} = \{E \subset X : E \text{ is linearly ordered}\}$ and order $\mathcal{F}$ by inclusion. If $\mathcal{M} \subset \mathcal{F}$ is linearly ordered, let $E = \cup \mathcal{M} = \bigcup_{A \in \mathcal{M}} A$. If $x, y \in E$ then $x \in A$ and $y \in B$ for some $A, B \subset \mathcal{M}$. Now $\mathcal{M}$ is linearly ordered by set inclusion so $A \subset B$ or $B \subset A$ i.e. $x, y \in A$ or $x, y \in B$. Sinse A and B are linearly order we must have either $x \leq y$ or $y \leq x$, that is to say E is linearly ordered. Hence by 3. there exists a maximal element $E \in \mathcal{F}$ which is the assertion in 2.

[53] Here is an easy proof if the elements of $\mathcal{F}$ happened to all be finite sets and there existed a set $P \in \mathcal{F}$ with a maximal number of elements. In this case the condition that $P \subset g(P)$ would imply that $P = g(P)$, otherwise $g(P)$ would have more elements than P.

construct $P_2 = \bigcup_{n=0}^{\infty} g^{(n)}(P_1) \in \mathcal{F}$ and $P_3 = \bigcup_{n=0}^{\infty} g^{(n)}(P_2) \in \mathcal{F}$, etc. etc. Then take $P_\infty = \cup_{n=0}^{\infty} P_n$ and start again with P_0 replaced by P_∞. Then keep going this way until eventually the sets stop increasing in size, in which case we have found our fixed point. The problem with this strategy is that we may never win. (This is very reminiscent of constructing measurable sets and the way out is to use measure theoretic like arguments.)

Let us now start the **formal proof.** Again let $P_0 \in \mathcal{F}$ and let $\mathcal{F}_1 = \{P \in \mathcal{F} : P_0 \subset P\}$. Notice that $\mathcal{F}_1$ has the following properties:

(1) $P_0 \in \mathcal{F}_1$.
(2) If $\Phi \subset \mathcal{F}_1$ is a totally ordered (by set inclusion) subset then $\cup\Phi \in \mathcal{F}_1$.
(3) If $P \in \mathcal{F}_1$ then $g(P) \in \mathcal{F}_1$.

Let us call a general subset $\mathcal{F}' \subset \mathcal{F}$ satisfying these three conditions a tower and let

$$\mathcal{F}_0 = \cap \{\mathcal{F}' : \mathcal{F}' \text{ is a tower}\}.$$

Standard arguments show that $\mathcal{F}_0$ is still a tower and clearly is the smallest tower containing P_0. (Morally speaking $\mathcal{F}_0$ consists of all of the sets we were trying to constructed in the "idea section" of the proof.)

We now claim that $\mathcal{F}_0$ is a linearly ordered subset of $\mathcal{F}$. To prove this let $\Gamma \subset \mathcal{F}_0$ be the linearly ordered set

$$\Gamma = \{C \in \mathcal{F}_0 : \text{ for all } A \in \mathcal{F}_0 \text{ either } A \subset C \text{ or } C \subset A\}.$$

Shortly we will show that $\Gamma \subset \mathcal{F}_0$ is a tower and hence that $\mathcal{F}_0 = \Gamma$. That is to say $\mathcal{F}_0$ is linearly ordered. Assuming this for the moment let us finish the proof. Let $P \equiv \cup\mathcal{F}_0$ which is in $\mathcal{F}_0$ by property 2 and is clearly the largest element in $\mathcal{F}_0$. By 3. it now follows that $P \subset g(P) \in \mathcal{F}_0$ and by maximality of P, we have $g(P) = P$, the desired fixed point. So to finish the proof, we must show that Γ is a tower.

First off it is clear that $P_0 \in \Gamma$ so in particular Γ is not empty. For each $C \in \Gamma$ let

$$\Phi_C := \{A \in \mathcal{F}_0 : \text{either } A \subset C \text{ or } g(C) \subset A\}.$$

We will begin by showing that $\Phi_C \subset \mathcal{F}_0$ is a tower and therefore that $\Phi_C = \mathcal{F}_0$.

1. $P_0 \in \Phi_C$ since $P_0 \subset C$ for all $C \in \Gamma \subset \mathcal{F}_0$. 2. If $\Phi \subset \Phi_C \subset \mathcal{F}_0$ is totally ordered by set inclusion, then $A_\Phi := \cup\Phi \in \mathcal{F}_0$. We must show $A_\Phi \in \Phi_C$, that is that $A_\Phi \subset C$ or $C \subset A_\Phi$. Now if $A \subset C$ for all $A \in \Phi$, then $A_\Phi \subset C$ and hence $A_\Phi \in \Phi_C$. On the other hand if there is some $A \in \Phi$ such that $g(C) \subset A$ then clearly $g(C) \subset A_\Phi$ and again $A_\Phi \in \Phi_C$.

3. Given $A \in \Phi_C$ we must show $g(A) \in \Phi_C$, i.e. that

$$g(A) \subset C \text{ or } g(C) \subset g(A). \tag{B.2}$$

There are three cases to consider: either $A \subsetneq C$, $A = C$, or $g(C) \subset A$. In the case $A = C$, $g(C) = g(A) \subset g(A)$ and if $g(C) \subset A$ then $g(C) \subset A \subset g(A)$ and Eq. (B.2) holds in either of these cases. So assume that $A \subsetneq C$. Since $C \in \Gamma$, either $g(A) \subset C$ (in which case we are done) or $C \subset g(A)$. Hence we may assume that

$$A \subsetneq C \subset g(A).$$

Now if C were a proper subset of $g(A)$ it would then follow that $g(A) \setminus A$ would consist of at least two points which contradicts the definition of g. Hence we must have $g(A) = C \subset C$ and again Eq. (B.2) holds, so Φ_C is a tower.

It is now easy to show Γ is a tower. It is again clear that $P_0 \in \Gamma$ and Property 2. may be checked for Γ in the same way as it was done for Φ_C above. For Property 3., if $C \in \Gamma$ we may use $\Phi_C = \mathcal{F}_0$ to conclude for all $A \in \mathcal{F}_0$, either $A \subset C \subset g(C)$ or $g(C) \subset A$, i.e. $g(C) \in \Gamma$. Thus Γ is a tower and we are done. ■