Axiom of Choice, Zorn's Lemma and the Well-ordering Principle

Let us briefly revisit the Axiom of Choice.

Proposition. *The following statements are equivalent:*

(AC) *for every family $\{A_i\}_{i\in I}$ of non-empty sets, the Cartesian product $\prod_{i\in I} A_i$ is non-empty; and*

(AC') *given a non-empty set X, there is a "choice function",*
$f : \mathcal{P}(X) \setminus \{\varnothing\} \to X$, for which $f(A) \in A$ for each A in $\mathcal{P}(X) \setminus \{\varnothing\}$.

Proof. (AC) $\Rightarrow$ (AC'). There exists $(x_A)_{A\in\mathcal{P}(X)\setminus\{\varnothing\}} \in \prod_{\mathcal{P}(X)\setminus\{\varnothing\}} A$. Let $f(A) = x_A$.

(AC') $\Rightarrow$ (AC). Let $X = \bigsqcup_{i\in I} A_i$. Let $f : \mathcal{P}(X) \setminus \{\varnothing\} \to X$ be any choice function. Then $(f(A_i))_{i\in I} \in \prod_{i\in I} A_i$. □

Definition/Notation. Given any non-empty set, S, a *binary relation* R is simply a subset of the Cartesian product $S \times S$. We tend to write "$s\,R\,t$" instead of "$(s,t) \in R$".

Definition. Let S be a non-empty set. A binary relation $R =$ " $\leq$ " on S is called a *partial ordering* if it satisfies, for s, t, u in S

(i)	$s \leq s$	(reflexivity)
(ii)	$s \leq t,\ t \leq u \Rightarrow s \leq u$	(transitivity)
(iii)	$s \leq t,\ t \leq s \Rightarrow s = t$	(anti-symmetry)

We call the pair $(S, \leq)$ a *partially ordered set.*

In $(S, \leq)$, a *chain* is any subset C for which any two elements are comparable, i.e. for any s, t in C, either $s \leq t$ or $t \leq s$. If S itself is a chain in $(S, \leq)$, we say that $\leq$ is a *total ordering* on S. If A is any subset of S, an *upper bound* for A is any u in S for which $s \leq u$ for s in A.

A *well-ordering* is any ordering $\leq$ on S such that in any non-empty subset A there is a *minimal element*, i.e. a in A such that $a \leq s$ for s in A.

Observe that a well-ordered set is totally ordered.

Theorem. *The following statements are equivalent:*

(i) Axiom of Choice (AC'): *for every non-empty X, there is a* choice function, *i.e.* $f : \mathcal{P}(X) \setminus \{\varnothing\} \to X$ *such that* $f(A) \in A$ *for each* A*.*

(ii) Hausdorff's Maximality Principle: *in any partially ordered set* $(S, \leq)$ *there is a* maximal chain, *i.e. a chain* M *for which no* $M \cup \{s\}$ *is a chain for any* s *in* $S \setminus M$*.*

(iii) Zorn's Lemma: *if in a partially ordered set* $(S, \leq)$*, each chain has an upper bound, then there is a* maximal element m *for* S*, i.e.* $m \leq s$ *implies* $m = s$*.*

(iv) Well-ordering Principle: *any non-empty set* S *admits a well-ordering.*

Proof. (i) $\Rightarrow$ (ii). We first consider an ancillary result, based on axiom of choice.

(I) *Consider the partially ordered set* $(\mathcal{P}(X), \subseteq)$*. Let* $\mathcal{F} \subseteq \mathcal{P}(X)$ *satisfy*

- $\varnothing \in \mathcal{F}$*, and*
- *if* $\mathcal{K} \subseteq \mathcal{F}$ *is a chain, then* $\bigcup_{K \in \mathcal{K}} K \in \mathcal{F}$*.*

Then $\mathcal{F}$ *contains an element* M *such that* $M \cup \{x\} \notin \mathcal{F}$ *for any* $x \in X \setminus M$*.*

Let us prove (I). For each A in $\mathcal{F}$ let

$$A^* = \{x \in X : A \cup \{x\} \in \mathcal{F}\}.$$

We use our assumption of (AC') to fix a

$$\text{choice function } f : \mathcal{P}(X) \setminus \{\varnothing\} \to X$$

and then let

$$F : \mathcal{F} \to \mathcal{P}(X), \quad F(A) = \begin{cases} A \cup \{f(A^*)\} & \text{if } A^* \neq \varnothing \\ A & \text{otherwise.} \end{cases}$$

We note that $f(A^*) \in A^*$ for each A in $\mathcal{F}$ for which $A^* \neq \varnothing$, and hence $F(A) \in \mathcal{F}$, by definition of A^*; i.e. $F : \mathcal{F} \to \mathcal{F}$.

We define an $(f, \mathcal{F})$-*tower* to be any subcollection $\mathcal{T} \subseteq \mathcal{F}$ for which

- $\varnothing \in \mathcal{T}$,
- $A \in \mathcal{T} \Rightarrow F(A) \in \mathcal{T}$
- if $\mathcal{K} \subseteq \mathcal{T}$ is a chain, then $\bigcup_{K \in \mathcal{K}} K \in \mathcal{T}$.

Notice that $\mathcal{F}$, itself, is an $(f, \mathcal{F})$-tower, and that the intersection of any family of $(f, \mathcal{F})$-towers is again an $(f, \mathcal{F})$-tower. Hence

$$\mathcal{T}_0 = \bigcap \{\mathcal{T} : \mathcal{T} \subseteq \mathcal{F} \text{ is an } (f, \mathcal{F})\text{-tower}\}$$

is an $(f, \mathcal{F})$-tower, in fact the minimal $(f, \mathcal{F})$-tower. Notice, $\varnothing \in \mathcal{T}_0$, so $\mathcal{T}_0$ is non-empty.

We aim to show that $(\mathcal{T}_0, \subseteq)$ is totally ordered. To this end, we call a set C in $\mathcal{T}_0$ *comparable* (in $\mathcal{T}_0$), if for A in $\mathcal{T}_0$, either $A \subseteq C$ or $C \subseteq A$. For such C consider the family

$$\mathcal{T}_C = \{A \in \mathcal{T}_0 : A \subsetneq C\} \cup \{C\} \cup \{A \in \mathcal{T}_0 : F(C) \subseteq A\}.$$

We observe that $\varnothing \in \mathcal{T}_C$. If $A \in \mathcal{T}_C$ then $f(A) \in \mathcal{T}_0$, and, using the assumption the C is comparable, we see that

- if $A \subsetneq C$, then $F(A) \subseteq C$, since otherwise, in the case that $A^* \neq \varnothing$, we would have $A \subsetneq C \subsetneq F(A) = A \cup \{f(A^*)\}$, which is clearly impossible; or
- if $A = C$ or if $F(C) \subseteq A$ then $C \subseteq A \subseteq F(A)$;

hence $f(A) \in \mathcal{T}_C$. Moreover, if $\mathcal{K}$ is a chain in $\mathcal{T}_C$, then let $B = \bigcup_{K \in \mathcal{K}} K$. Indeed if each $K \subseteq C$, then $B \subseteq C$; and if $F(C) \subseteq K$ for some K, then $F(C) \subseteq B$. Thus $\mathcal{T}_C$ is a tower, in which case we must have $\mathcal{T}_C = \mathcal{T}_0$, as $\mathcal{T}_0$ is the minimal tower in $\mathcal{F}$.

It follows that $F(C)$ is comparable in $\mathcal{T}_0$ if C is. Thus the family of comparable sets, $\mathcal{C}$, satisfies the first two axioms of an $(f, \mathcal{F})$-tower; it remains to check the third. If $\mathcal{K}$ is a chain in $\mathcal{C}$, let $B = \bigcup_{K \in \mathcal{K}} K$. If $A \in \mathcal{T}_0$ then either $A \subseteq K$ for some K, in which case $A \subseteq B$; or $K \subseteq A$ for all K, in which case $B \subseteq A$. Thus $B \in \mathcal{C}$. Hence $\mathcal{C}$ is itself an $(f, \mathcal{F})$-tower, and again by minimality of $\mathcal{T}_0$, we see that $\mathcal{C} = \mathcal{T}_0$. Hence we have that $(\mathcal{T}_0, \subseteq)$ is indeed totally ordered, hence a chain in $(\mathcal{F}, \subseteq)$.

Now we let $M = \bigcup_{T \in \mathcal{T}_0} T \in \mathcal{T}_0$. If it were the case that $M^* \neq \varnothing$, we would have that $F(M) = M \cup \{f(M^*)\} \in \mathcal{T}_0$ since $\mathcal{T}_0$ is a tower. But this violates the fact that $f(M^*) \notin M$. Hence $M^* = \varnothing$ which proves (I).

(II) We now use (I) to prove (ii). Given a partially ordered set $(S, \leq)$, let $\mathcal{F}$ denote the set of all chains in S. We remark that $\varnothing$ is trivially a chain. Any chain $\mathcal{K}$ in $(\mathcal{F}, \subseteq)$ has that $C = \bigcup_{K \in \mathcal{K}} K$ is a chain, i.e. any two elements of C must live in some K, and hence are comparable. Any M, arising form the conclusion of (I), is a maximal chain.

(ii) $\Rightarrow$ (iii). Suppose $(S, \leq)$ is a partially ordered set in which each chain has a maximal element. Let M be a maximal chain in $(S, \leq)$ and m be an upper bound for M. Then $M \cup \{m\}$ is a chain, and hence equal to M by maximality of M, i.e. $m \in M$. Moreover, if any s in S satisfies $m \leq s$, then $M \cup \{s\}$ is a chain, from which it again follows that $s \in M$, hence $s \leq m$. But then $s = m$, so m is a maximal element.

(iii) $\Rightarrow$ (iv). Let

$$\mathcal{W} = \{(A, R_A) : A \in \mathcal{P}(X),\ R_A \subseteq A \times A \text{ is a well-ordering}\}.$$

(Notice we use R_A instead of symbol $\leq_A$.) We note that $\mathcal{W} \neq \varnothing$. Indeed, any countable set $A \subseteq X$ can be well-orderd. We define a relation on $\mathcal{W}$ by

$$(A, R_A) \preceq (B, R_B) \;\Leftrightarrow\; (A, R_A) \text{ is an } \textit{initial segment} \text{ of } (B, R_B)$$

which is to say that

- $A \subseteq B$,
- $R_B \cap (A \times A) = R_A$, and
- for a in A and b in $B \setminus A$, we have aR_Bb.

Then $\preceq$ is evidently reflexive, transitive and anti-symmetric, hence a partial order on $\mathcal{W}$.

Let $\mathcal{C}$ be a chain in $(\mathcal{W}, \preceq)$. Let $U = \bigcup_{(C, R_C) \in \mathcal{M}} C$ and for s, t in U, let sR_Ut whenever $s, t \in C$ with sR_Ct, for some $(C, R_C) \in \mathcal{C}$. Then R_U is trivially well-defined. If $A \subseteq U$ is non-empty, there is some (C, R_C) in $\mathcal{C}$ for which $A \cap C \neq \varnothing$, and thus admits a minimal element a_C. Observe that if $A \cap C' \neq \varnothing$ for another $(C', R_{C'})$ in $\mathcal{C}$, then $C \subseteq C'$, say, and we see that $a_{C'} = a_C$, since

(C, R_C) is an initial segment of $(C', R_{C'})$. Hence, $(U, R_U) \in \mathcal{W}$ and is an upper bound for $\mathcal{C}$.

Hence, by Zorn's lemma, $\mathcal{W}$ admits a maximal element (M, R_M). If there were s in $S \setminus M$, we could let $M' = M \cup \{s\}$ and extend R_M to M' by assigning $tR_{M'}s$ for all t in M. But then $(M', R_{M'}) \in \mathcal{W}$, which would violate the maximlity of (M, R_M). Hence R_M is a well-ordering on S.

(iv) $\Rightarrow$ (i). Suppose $\leq$ is a well-ordering on X. Let $f(A)$ be the minimal element of A for each A in $\mathcal{P}(X) \setminus \{\varnothing\}$. □