The axiom of choice and equivalent statements 21/7/2010
Takis Konstantopoulos

Abstract

The following is a more or less standard treatment of some of the equivalent statements to one of the most fundamentamental tenet of Set Theory and, indeed, all of Mathematics, namely, the Axiom of Choice.

1 The basic axiom

THE AXIOM OF CHOICE: For every collection $\{A_t, t \in T\}$ of sets there is a choice function $g : T \to \cup_{t\in T} A_t$ such that $g(t) \in A_t$ for all $t \in T$.

Ever since its conception, the Axiom of Choice has been proven an indispensable pillar of Mathematics. Without it, lots of the things, as we know them, wouldn't hold. One can, for instance, construct examples of sets in 3-dimensional Euclidean space which have no well-defined volume (they have no volume at all!) precisely because of the Axiom of Choice. If we drop the axiom then we get into all kinds of trouble, so we accept it. On the other hand, there are, from time to time, efforts to see whether one can do without it. For instance, Doyle and Conway [1] show that if A and B are sets such that $\{1, 2, 3\} \times A$ and $\{1, 2, 3\} \times B$ have the same cardinality, then so do A and B (by two sets having the same cardinality we mean that there is a bijection between them); and they do so, without the Axiom of Choice, but with a lot of effort! In what follows, we explain carefully why the Axiom Choice can be expressed in other, perhaps more useful, manners. There is nothing new here. This is not a research paper.

2 Equivalent statements

A collection of sets is said to be of finite character if membership of a set in the collection is determined by membership of each of its finite subsets.

Trivially, the empty set $\varnothing$ belongs to any collection of sets of finite character.

For example, the collection of sets in $\mathbb{R}^d$ which have diameter at most 1 is of finite character.

We wish to prove that:

JOHN TUKEY'S LEMMA: If a collection of sets is of finite character then it necessarily contains maximal (in the sense of inclusion) sets.

Suppose $\mathscr{F}$ is a collection of sets of finite character with no maximal elements. This means that for every $\mathscr{F}$-set there is a larger $\mathscr{F}$-set. We assume that the axiom of choice holds. Then for each $\mathscr{F}$-set F we can pick a strictly larger $\mathscr{F}$-set $f(F)$. That is, we have a rule f which can be described as "pick a bigger set".

By the assumption that there are no maximal elements, we can apply the rule f as many times as we wish and we will still remain within $\mathscr{F}$:

$$F \mapsto f(F) \mapsto f(f(F)) \mapsto f(f(f(F))) \mapsto \cdots$$

So, starting with any set F we can create a chain

$$\{f^{(n)}(F), n = 0, 1, \ldots\}$$

Note that this chain does not contain $f^{(\infty)}(F) := \cup_n f^{(n)}(F)$ because this set cannot be obtained by applying f finitely many times. But we can just add this set to the chain and get a strictly bigger chain. In fact we can keep going:

$$\{f^{(n)}(F), n = 0, 1, \ldots\} \cup \{f^{(\infty)}(F), f(f^{(\infty)}(F)), \ldots\};$$

and we can go further by adding $\cup_n f^{(n)}(f^{(\infty)}(F))$; and we can keep going and going and going, just as in the construction of ordinal numbers.

This prompts us to define a family $\mathscr{I}$ of sets which is like the above. Such a family has the following properties: 1) $\varnothing \in \mathscr{I}$;
2) if $F \in \mathscr{I}$ then $f(F) \in \mathscr{I}$;
3) it is closed under increasing unions.
We will not insist that $\mathscr{I}$ be a chain; only that it satisfies the above properties. Of course, a closed chain will satisfy 1) and 2) and 3), but there may be other families that also do so. Let us call such a family and f-family.

It is important to notice that
the intersection of f-families is also an f-family.

Hence we can define

$$\mathscr{I}_0 := \cap\{\mathscr{I} : \mathscr{I} \text{ is an } f\text{-family}\}$$

and be sure that $\mathscr{I}_0$ is an f-family.

We next have the hunch that
$\mathscr{I}_0$ should be a chain.

If we manage to prove our hunch we will have reached a contradiction because:
On one hand, the set $I^* := \cup\{I : I \in \mathscr{I}_0\}$ is a member of $\mathscr{I}_0$ because $\mathscr{I}_0$ is an f-family and a chain.
On the other hand, we can apply the rule f to I^* and obtain a set $f(I^*)$ which is simultaneously strictly bigger than I^* but also a member of $\mathscr{I}_0$ since $\mathscr{I}_0$ is an f-family. And this implies the absurdity that I^* is a strict subset of itself!

So now we know what remains to be proved: Our hunch!

In other words, we wish to prove that
If $A, C \in \mathscr{I}_0$ then either $C \subset A$ or $A \subset C$.

But let us not forget that $\mathscr{I}_0$ is an f-family which means, in particular, that it is closed under increasing unions. If so, we have a stronger hunch:
If $A, C \in \mathscr{I}_0$ then either $C \subset A$ or $f(A) \subset C$.

If we prove the stronger hunch then we are done.

But let us take things one at a time. Maybe we can't prove the stronger hunch for all sets in $\mathscr{I}_0$. Perhaps there is a subclass of $\mathscr{I}_0$ for which if A belongs to this subclass and $C \in \mathscr{I}_0$ then either $C \subset A$ or $f(A) \subset C$. Here is a candidate for this subclass: the set A must be an f-limit point. That is, if $B \neq A$, $B \subset A$ then $f(B) \subset A$. Think about it: If A is an f-limit point and $\mathscr{I}_0$ is a chain then the stronger hunch out to be true.

We have a definition and two claims:

<u>Definition:</u> A set is an f-limit point if for every set $B \neq A$, $B \subset A$ we also have $f(B) \subset A$. Let LIM_f be the set of f-limit points.

<u>Claim 1:</u> If $A \in \mathrm{LIM}_f$ and $C \in \mathscr{I}_0$ **then either** $C \subset A$ **or** $f(A) \subset C$.

<u>Claim 2:</u> All sets in $\mathscr{I}_0$ are f-limits.

PROOF OF CLAIM 1: Claim 1 is equivalent to proving that if $A \in \mathrm{LIM}_f$ then

$$\{C \in \mathscr{I}_0 : \ C \subset A \text{ or } f(A) \subset C\}$$

is an f-family. If so, then it will be equal to $\mathscr{I}_0$ for the latter is the smallest f-family. Let C be an element of this family. We want to prove that $f(C)$ is also an element, i.e. that $f(C) \subset A$ or $f(A) \subset f(C)$. Since C is an element of this family, either (i) $C \subsetneq A$ or (ii) $C = A$ or (iii) $f(A) \subset C$. If (i) then $f(C) \subset A$ because $A \in \mathrm{LIM}_f$. If (ii) then $f(A) = f(C)$, i.e. $f(A) \subset f(C)$. If (iii) then $f(A) \subset C \subsetneq f(C)$. In either case, $f(C) \subset A$ or $f(A) \subset f(C)$, which means that $f(C)$ is also a member of the above family.
Let now $(C_t, t \in T)$ be a chain of sets of the above family. Let $C^* := \cup_{t \in T} C_t$. Every C_t is a member of the above family, that is, for each $t \in T$, either $C_t \subset A$ or $f(A) \subset C_t$. There are two alternatives: either $f(A) \subset C_{t_0}$ for some $t_0 \in T$, in which case $f(A) \subset C^*$, or there is no such t_0, in which case $C_t \subset A$ for all $t \in T$ and, consequently, $C^* \subset A$. Thus C^* is a member of the above family, and this concludes the proof of claim 1. □

PROOF OF CLAIM 2: Our goal to prove that $\mathscr{I}_0 \subset \mathrm{LIM}_f$. Our goal will be achieved if we prove that LIM_f is an f-family. So here we go again, there are two things to be proved.
First, let $A \in \mathrm{LIM}_f$ and show that $f(A) \in \mathrm{LIM}_f$.
Second, show that LIM_f is closed under increasing unions.
To prove the first, pick an $A \in \mathrm{LIM}_f$. Let $B \subsetneq f(A)$ and prove that $f(B) \subset f(A)$. By Claim 1, either $B \subset A$ or $f(A) \subset B$. But the second alternative is incompatible with the choice of B as a strict subset of $f(A)$. Hence $B \subset A$. Since A is an f-limit, $f(B) \subset A$ and so $f(B) \subset f(A)$.
To prove the second, let $(A_t, t \in T)$ be a chain of f-limit sets and show that $A^* := \cup_{t \in T} A_t$ is also an f-limit: Let $B \subsetneq A^*$ and show that $f(B) \subset A^*$. Since each A_t is f-limit, we have, by Claim 1, for each $t \in T$, $B \subset A_t$ or $f(A_t) \subset B$. There are two alternatives: either $B \subset A_{t_0}$ for some $t_0 \in T$, or there is no such t_0, in which case $f(A_t) \subset B$ for all $t \in T$. But the latter case leads to the immediate absurdity $B \supset \cup_{t \in T} f(A_t) \supsetneq \cup_{t \in T} A_t = A^* \supsetneq B$. So we are left with the possibility of the existence of a t_0 with $B \subset A_{t_0}$. If B is a strict subset of A_{t_0} then (the latter being an f-limit) $f(B) \subset A_{t_0} \subset A^*$. If $B = A_{t_0}$ then B itself is an f-limit and, by Claim 1, either $A^* \subset B$ (which is impossible because $B \subsetneq A^*$) or $f(B) \subset A^*$, which is the only possibility. Hence, even when $B = A_{t_0}$ for some t_0, we have $f(B) \subset A^*$. We are done. □

Let us now finish the proof of Tukey's lemma. We have shown that our stronger hunch is correct. So, in particular, $\mathscr{I}_0$ is a chain. Let $M := \cup\{A : \ A \in \mathscr{I}_0\}$. Since $\mathscr{I}_0$ is an f-family, $M \in \mathscr{I}_0$. Again, since $\mathscr{I}_0$ is an f-family, $f(M) \in \mathscr{I}_0$ and $f(f(M)) \in \mathscr{I}_0$, etc. But this can't be because M is a strict subset of $f(M)$. □

The next thing to be proved is:

FELIX HAUSDORFF'S PRINCIPLE: Every poset has maximal chains.

Let $(P, \leq)$ be a poset. Let $\mathfrak{C}$ be the collection of all chains in this poset. Then $\mathfrak{C}$ has finite character: $\mathscr{C} \in \mathfrak{C}$ iff every two element subset of $\mathscr{C}$ is a chain. Hence $\mathfrak{C}$ has maximal elements. □

And here is a further fact:

MAX AUGUST ZORN'S LEMMA: If every chain in a poset has an upper bound then the poset has maximal elements.

Let $(P, \leq)$ be a poset. By Hausdorff's principle, there is a maximal chain $\mathscr{C}$. By assumption, $\mathscr{C}$ has an upper bound $m \in P$. This upper bound must be a maximal element of the poset, otherwise we would be able to create an even larger chain $\mathscr{C} \cup \{m\}$, contradicting the maximality of $\mathscr{C}$. □

Let us now define the notion of a well-ordered set as a linearly ordered set such that every subset contains a minimum. Perhaps the most surprising of all is the following statement:

ERNST ZERMELO'S THEOREM: Every set can be well-ordered.

Think about it: The real numbers can be well-ordered. But how? Certainly, the usual ordering is not a well-ordering. Well, the answer is that we do not know how we can do it, but we can prove that it can be done.

Here is a sketch of proof (details can be filled in):

Let S be a set. Let $\mathscr{X}$ be the collection of all well-ordered sets $(W, \leq)$ with $W \subset S$. The idea is that $\mathscr{X}$ can be partially ordered naturally by:

$$(W_1, \leq_1) \preceq (W_2, \leq_2) \iff W_1 \subset W_2 \text{ and } \leq_2 \text{ extends } \leq_1 .$$

Hence $(\mathscr{X}, \preceq)$ is a poset. We want to show that it has a maximal element (which will turn out to be a well-ordering on S itself). To show this, it suffices to show that every chain has an upper bound (this is Zorn's lemma). So let $((W_t, \leq_t), t \in T)$ be a chain in $\mathscr{X}$. There is a natural candidate for an upper bound of this chain, namely,

$$W := \cup_{t \in T} W_t, \quad \leq := \cup_{t \in T} \leq_t .$$

It takes a bit to show that, indeed, $(W, \leq)$ is an upper bound for $((W_t, \leq_t), t \in T)$, and that $\leq$ is a *linear ordering* on W. Applying Zorn's lemma, we have the existence of a maximal element $(W^*, \leq^*)$ for $\mathscr{X}$. By the preceding, $\leq^*$ is a linear ordering on W^*. By maximality, $W^* = S$, and we are done. □

So, starting from the Axiom of Choice, we proved a sequence of four, seemingly stronger, statements. But they are not stronger. In fact, they are all equivalent to the Axiom of Choice because:

Zermelo's theorem implies the axiom of choice.

Indeed, if $\{A_t, t \in T\}$ is a collection of sets, then $A := \cup_{t \in T} A_t$ can be well-ordered and so each subset has a minimum. In particular, for each t, $A_t \subset A$ and so A_t has a minimum. So here is then the choice function:

$$t \mapsto \text{ minimum of } A_t.$$

□

References

[1] Doyle, P.G. and Conway, J.H. (1994). Division by three. http://www.math.dartmouth.edu/~doyle/docs/three/three.pdf. http://arxiv.org/PS_cache/math/pdf/0605/0605779v1.pdf/