Products
and
The Axiom of Choice

John McCuan

January 10, 2020

A cursory scan of Gunning's book *An Introduction to Analysis* suggests that he has not included any discussion of general Cartesian products nor any mention of the axiom of choice. Bartle mentions the axiom of choice, but doesn't state it either, so I assume this can be forgiven. On the other hand, these topics are interesting, useful, and not that difficult to say some things about, so let's include them.

As extra motivation, I stated an exercise in class:

> Find a surjective function $f : A \to B$ such that there is no function $g : B \to A$ for which $f \circ g = \mathrm{id}_B$.

I can't say the indignation was quite universal, but there was a healthy movement lobbying for the replacement of this exercise with a contrarian claim:

> Claim: Given any surjective function $f : A \to B$, there **is** a function $g : B \to A$ for which $f \circ g = \mathrm{id}_B$.

"Fair enough," I said, but you need to prove it. And that wasn't so easy. I don't think it can be proved without the Axiom of Choice. At least I don't see a way to prove it without the Axiom of Choice, and I think it's more or less equivalent. I'll talk about some other unsettling claims that are equivalent to the Axiom of Choice below, and you can decide whether you want to opt for the contrarian claim, my exercise, or neither.

1 Index Sets

When we have a surjective function $f : A \to B$, then we can think of A as an index set for the set B. Notice that this doesn't mean the elements of B are uniquely indexed, because we're not saying anything about f being an injective function. Nevertheless, the point is that we can think of elements of B as **indexed**, though not uniquely, by elements of A, and write

$$b_a = f(a).$$

This is not saying anything profound. We're basically just suppressing the name of the function f. It may be recalled that the definition of a function we gave:

> Given two sets A and B, a **function** f from A to B is a rule or correspondence with assigns to each $x \in A$ and unique $y \in B$.

(probably) had the name of the function in it. Or else, we quickly gave the function a name and wrote $f : A \to B$ and $y = f(x)$, and a host of other things involving the name like $f(E) = \{f(x) : x \in E\}$. But the *idea of a function* does not need the name. We can also state the definition without it:

> Given two sets A and B, a **function** from A to B is a rule or correspondence with assigns to each $x \in A$ and unique $y \in B$.

The *idea of a function* still survives (quite happily) even without the name. This is a little bit like roses.[1] Thus, we are *thinking of the function*

$$a \mapsto b_a.$$

This notion should be especially familiar from sequences. A sequence is a surjective function from $\mathbb{N}$ or $\mathbb{N}_0$ to some set. Such a function, often denoted by

$$\{y_j\}_{j\in\mathbb{N}} = \{y_1, y_2, y_2, \ldots\},$$

can be very non-injective:

$$\{1, 1, 1, 1, 1, \ldots\}.$$

What might be a little new here is using other (and especially bigger) indexing sets like, for example, $\mathbb{R}$. The set of all open lines $\{\ell_m\}_{m\in\mathbb{R}}$ where

$$\ell_m = \{(x, mx) \in \mathbb{R}^2 : x \in \mathbb{R}\}$$

is a good example.

[1] Act II, Scene II of *Romeo and Juliet* by William Shakespeare.

2 Products

You may remember the fabled Axiom of Unions from set theory. According to this axiom, whenever you have a collection of sets $\{A_\alpha\}_{\alpha\in\Gamma}$ indexed by Γ, there exists a set (called the **union**) containing all the elements in all the sets A_α. This set is denoted by

$$\bigcup_{\alpha\in\Gamma} A_\alpha$$

or sometimes

$$\bigcup\{A_\alpha : \alpha \in \Gamma\},$$

but I like the first one better than the second one. Okay, now it's time to pay attention:

> The (Cartesian) **product** of a collection of sets $\{A_\alpha\}_{\alpha\in\Gamma}$ is the set of all functions $p : \Gamma \to \cup A_\alpha$ with the property that
>
> $$p(\alpha) \in A_\alpha.$$
>
> The product is denoted by $\prod_{\alpha\in\Gamma} A_\alpha$ and elements in the product are denoted by $(a_\alpha)_{\alpha\in\Gamma}$.

It's fine to think of

$$p = (a_\alpha)_{\alpha\in\Gamma} \in \prod_{\alpha\in\Gamma} A_\alpha \qquad \text{where} \qquad p(\alpha) = a_\alpha$$

as a Γ-tuple. The only thing that's different from the "ordered pairs" and "ordered triples" and so on you've seen before, is that we haven't said anything about an ordering on the indexing set Γ—and there might not be one. So thinking about ordered Γ-tuples doesn't really make sense. They're just Γ-tuples. The product is the set of all Γ-tuples.

3 The Axiom of Choice

If one of the sets in a product is the empty set, then the product is also the empty set. If both sets in a Cartesian product of two sets are nonempty, then the Cartesian product itself is nonempty. To see this, say A_1 and A_2 are nonempty. Then there are elements $x_1 \in A_1$ and $x_2 \in A_2$. Then $(x_1, x_2) \in A_1 \times A_2 = \prod_{j=1}^{2} A_j$. And this kind of assertion follows by induction for any **finite product**

$$\prod_{j=1}^{k} A_j.$$

The Axiom of Choice is the following:

> The product of a nonempty collection of nonempty sets is nonempty.

That seems reasonable.

On the other hand, (we haven't done it, but) one can construct the real numbers $\mathbb{R}$ without using this choice Axiom, construct the finite product $\mathbb{R}^3 = \mathbb{R} \times \mathbb{R} \times \mathbb{R}$, and then prove the following thing using the Axiom of Choice:

Theorem 1 (Banach-Tarski paradox) *There are disjoint sets A_1, A_2, A_3, A_4 and A_5 in*

$$B_1(\mathbf{0}) = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 < 1\}$$

and **rigid motions**[2] *of space f_1, f_2, f_3, f_4 and f_5 such that*

$$\bigcup_{j=1}^{5} f_j(A_j) = B_2(0) = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 < 2\}.$$

This says you can disassemble a unit size ball in $\mathbb{R}^3$ into five pieces, and then rearrange the pieces into a ball (the whole entire thing) that has twice the radius.

Another weird consequence of the Axiom of Choice

Here is something else: Let Σ be a subset of the collection of all subsets of the unit interval $(0, 1)$ in the real numbers. That is, $\Sigma \subset \mathcal{P}(0, 1)$. Assume that Σ contains all the intervals[3] and has the following three closure properties:

1. $A \cup B \in \Sigma$ whenever $A, B \in \Sigma$, and

[2] Which are special bijections $f : \mathbb{R}^3 \to \mathbb{R}^3$ consisting of rotations and translations. We haven't defined or discussed these functions, but their existence and properties should be intuitively clear. In short, you know these functions.

[3] We haven't defined intervals (or real numbers) yet, but again you can imagine that making such definitions is possible.

2. $A \backslash B \in \Sigma$ whenever $A, B \in \Sigma$.

3. $\{x + t : x \in A \text{ and } x + t < 1\} \cup x + t - 1 : x \in A \text{ and } x + t \geq 1\}$ whenever $A \in \Sigma$. The set defined here is called the **translation of A (modulo 1)**.

In particular, the second property implies the empty set $\phi \in \Sigma$.

A function μ which assigns non-negative numbers to sets in Σ, i.e., a function $\mu : \Sigma \to [0, 1]$, is a **Euclidean pre-measure** if the following properties hold:

1. $\mu(\phi) = 0$.

2. The measure of an interval is its length. So $\mu[a, b] = \mu[a, b) = \mu(a, b] = b - a$ whenever $0 < a \leq b < 1$.

3. If A and B are in Σ and $A \cap B = \phi$, then

$$\mu(A \cup B) = \mu(A) + \mu(B).$$

   This is called **finite additivity**.

4. If $\{A_j\}_{j=1}^{\infty}$ is a countable collection of sets in Σ for which $A_i \cap A_j = \phi$ for $i \neq j$, and it happens to be the case that the union $\cup A_j \in \Sigma$, then

$$\mu\left(\bigcup_{j=1}^{\infty} A_j\right) = \sum_{j=1}^{\infty} \mu(A_j).$$

   This is called **countable additivity**.

5. If $A \in \Sigma$ and $0 < t < 1$, then

$$B = \{x + t : x \in A \text{ and } x + t < 1\} \cup x + t - 1 : x \in A \text{ and } x + t \geq 1\}$$

   satisfies $B \in \Sigma$ and $\mu(B) = \mu(A)$. Rememebeer this set B is the translation of A (modulo 1), and this condition is called **translation invariance**.

Note that, in particular, one must have $(0, 1) \in \Sigma$ and $\mu(0, 1) = 1$. Therefore, if we take any subset $E \subset (0, 1)$ and look at countable unions of sets $A_1, A_2, A_3, \ldots$ in Σ for which $E \subset \cup A_j$, then we can always take $A_1 = (0, 1)$ and $A_2 = A_3 = A_4 = \cdots = \phi$ and have $\sum m(A_j) = 1$. This means the set of **extended real numbers**[4]

$$M(E) = \left\{ \sum_{j+1}^{\infty} \mu(A_j) : (A_j) \in \prod_{j=1}^{\infty} \Sigma \text{ and } E \subset \cup A_j \right\} \tag{1}$$

[4]That is, numbers in $[0, \infty]$ with the arithmetic $a + \infty = \infty$ for $a \in \mathbb{R}$.

always contains the number 1 and is bounded below by 0. The **completeness** of the real numbers, which we will cover in detail soon, says that a nonempty set of real numbers A which is bounded below has a **greatest lower bound** in the real numbers. this greatest lower bound (when it exists) is called the **infemum** of the set A and is denoted $\inf A$.

Theorem 2 (existence of a nonmeasurable set) *Given any Euclidean pre-measure μ, and setting*

$$m(E) = \inf(M(E) \cap \mathbb{R}), \tag{2}$$

there exist sets $N, E \subset (0, 1)$ such that

$$m(E \cap N) + m(E \backslash N) \neq m(E). \tag{3}$$

The set N is called a **nonmeasurable set**.

This result requires perhaps a little explanation to understand why it is so strange. Hopefully the five conditions defining a Euclidean premeasure should present themselves as

> *conditions satisfied by any reasonable notion of a measure defined on subsets of $(0, 1)$.*

You should consider each of these conditions individually and decide if you think it is reasonable as a notion of length that extends to sets beyond just intervals. Of course, a place to start with your deliberations is with sets that are simply disjoint unions of intervals.

Now the question is:

> *Can you extend a premeasure in some reasonable way to include more sets than just those in Σ?*

In particular, hiding in the background is the idea that you should at least be able to extend to a set which is large enough to be **closed under countable unions**, and you should still be able to ensure countable additivity and translation invariance. Notice that closure under countable unions is not required of the set Σ. Well, it turns out you can do that as follows:

1. The extension is given by m as defined in (2).

2. The larger set $\mathcal{C} \supset \Sigma$ is given by

$$\mathcal{C} = \{S \subset (0,1) : m(E \cap S) + m(E \backslash S) = m(E) \text{ for every } E \subset (0,1)\}. \qquad (4)$$

These are called the **Carathéodory measurable sets** and the condition

$$m(E \cap S) + m(E \backslash S) = m(E) \text{ for every } E \subset (0,1) \qquad (5)$$

on the set S is called the **Carathéodory measurability condition**.

So there are some things to prove here. First note that m is defined on **all** subsets of $(0,1)$, that is, the entire power set. So what Carathéodory is saying is to restrict m to the collection of measurable sets $\mathcal{C}$. Then one needs to prove:

1. $\Sigma \subset \mathcal{C}$.

2. $m(A) = \mu(A)$ when $A \in \Sigma$.

3. $\mathcal{C}$ satisfies the three closure properties satisfied by Σ.

4. $\mathcal{C}$ satisfies the additional closure property

$$\cup A_j \in \mathcal{C} \qquad \text{whenever } A_1, A_2, A_3, \ldots \in \mathcal{C}.$$

   This is called **closure under countable unions**. These four closure properties make $\mathcal{C}$ a **σ-algebra**. (You read this "sigma algebra.")

5. The function $m : \mathcal{C} \to [0,1]$ is countably additive and translation invariant. That is, m restricted to $\mathcal{C}$ is a translation invariant **measure**. Note that $m : \mathcal{C} \to [0,1]$ is called a measure and the function $m : \mathcal{P}(0,1) \to [0,1]$ is called the corresponding **outer measure**.

Carathéodory was able to prove all these things. What is more, is that this is essentially the *only way* you can get an extension. More precisely, if you have any extension m of μ to a σ-algebra containing Σ and m is a (countably additive) translation invariant measure, then the sets in your sigma algebra will satisfy (5); they will be Carathéodory measurable.

Now let's look finally at the Carathéodory measurability condition. What is says is that a measurable set S is one that "cuts every other set well with respect to m," in the sense that when you partition an arbitrary set E (not necessarily measurable) into two disjoint parts $E \cap S$ and $E \backslash S$, then the (outer) measures of the pieces add up to the (outer) measure of the set E.

Of course, maybe one can be skeptical concerning the measurement of sets that are much more complicated than intervals and take the view that there's no reason all sets should "cut well." On the face of it, however, it is rather surprising that one cannot find a translation invariant measure on the sigma algebra of **all** subsets of $(0, 1)$ having the property that the measure of an interval is its length. The existence of a nonmeasurable set says you can't do it—if you believe/adopt the Axiom of Choice. In fact, the existence of a nonmeasurable set is **equivalent** to the Axiom of Choice.

4 Surjective Functions and One-sided Inverses

Returning to our surjective function $f : A \to B$, the surjectivity condition implies the family of sets

$$f^{-1}(\{b\}) = \{x \in A : f(x) = b\} \tag{6}$$

is a family of nonempty sets indexed by the elements $b \in B$. Actually, this is not quite correct. There is another case to consider. This is the case in which $B = \phi$. In that case, the inverse images in (6) do not exist because there are no elements $b \in B$. In this case also we must have $A = \phi$, or else f cannot be a function, and then f must be what is called the **empty function** as well. In this case, the claim is pretty easy to prove. We can let $g : B \to A$ be the empty function. Then, of course, $f \circ g$ is the empty function, and we need only ask (and verify) that this is the identity function on $B = \phi$. I suppose the identity on the empty set must be the empty function. As mentioned before, that's the only function with target $B = \phi$, and we can say with certainty (and a certain level of opacity) that

$$\mathrm{id}_\phi = \phi : \phi \to \phi$$

is a function having the property that $\phi(b) = b$ for every $b \in \phi$.

Now then, if $B \neq \phi$, then the sets given in (6) indexed by B are all nonempty, and they are a nonempty collection. Thus, the Axiom of Choice says there is a **choice function**, i.e., an element

$$(a_b) \in \prod_{b \in B} f^{-1}(\{b\}).$$

The function p chooses an element $p(b) = a_b \in f^{-1}(\{b\})$. Thus, we may define $g : B \to A$ by $g(b) = p(b) = a_b$. And,

$$f \circ g(b) = f(a_p) = b \qquad \text{since} \qquad a_p \in f^{-1}(\{b\}).$$

That is, the choice function is precisely the function we need. And the claim is established (by the Axiom of Choice). $\square$

I see no other way to establish this claim.

Exercise 1 *If for every surjective function $f : A \to B$, there exists a function $g : B \to A$ such that $f \circ g = \mathrm{id}_B$, then the Axiom of Choice holds.*