Chapter udf

Choice

choice.1 Introduction

sth:choice:intro: sec

In **??**–**??**, we developed a theory of cardinals by treating cardinals as ordinals. That approach depends upon the Axiom of Well-Ordering. It turns out that Well-Ordering is equivalent to another principle—the Axiom of Choice—and there has been serious philosophical discussion of its acceptability. Our question for this chapter are: How is the Axiom used, and can it be justified?

choice.2 The Tarski-Scott Trick

sth:choice:tarskiscott: sec

In **??**, we defined cardinals as ordinals. To do this, we assumed the Axiom of Well-Ordering. We did this, for no other reason than that it is the "industry standard".

Before we discuss any of the philosophical issues surrounding Well-Ordering, then, it is important to be clear that we *can* depart from the industry standard, and develop a theory of cardinals *without* assuming Well-Ordering. We can still employ the definitions of $A \approx B$, $A \preceq B$ and $A \prec B$, as they appeared in **??**. We will just need a new notion of *cardinal*.

A naïve thought would be to attempt to define A's cardinality thus:

$$\{x : A \approx x\}.$$

You might want to compare this with Frege's definition of $\#xFx$, sketched at the very end of **??**. And, for reasons we gestured at there, this definition fails. Any singleton set is equinumerous with $\{\emptyset\}$. But new singleton sets are formed at every successor stage of the hierarchy (just consider the singleton of the previous stage). So $\{x : A \approx x\}$ does not exist, since it cannot have a rank.

To get around this problem, we use a trick due to Tarski and Scott:

Definition choice.1 (Tarski-Scott). For any formula $\varphi(x)$,[1] let $[x : \varphi(x)]$ be the set of all x, of least possible rank, such that $\varphi(x)$ (or $\emptyset$, if there are no φs).

We should check that this definition is legitimate. Working in **ZF**, **??** guarantees that $\text{rank}(x)$ exists for every x. Now, if there are any entities satisfying φ, then we can let α be the least rank such that $(\exists x \subseteq V_\alpha)\varphi(x)$, i.e., $(\forall \beta \in \alpha)(\forall x \subseteq V_\beta)\neg\varphi(x)$. We can then define $[x : \varphi(x)]$ by Separation as $\{x \in V_{\alpha+1} : \varphi(x)\}$.

Having justified the Tarski-Scott trick, we can now use it to define a notion of cardinality:

Definition choice.2. The TS-cardinality of A is $\text{tsc}(A) = [x : A \approx x]$.

The definition of a TS-cardinal does not use Well-Ordering. But, even without that Axiom, we can show that TS-*cardinals* behave rather like *cardinals* as defined in **??**. For example, if we restate **??** and **??** in terms of TS-cardinals, the proofs go through just fine in **ZF**, without assuming Well-Ordering.

Whilst we are on the topic, it is worth noting that we can also develop a theory of ordinals using the Tarski-Scott trick. Where $\langle A, <\rangle$ is a well-ordering, let $\text{tso}(A, <) = [\langle X, R\rangle : \langle A, <\rangle \cong \langle X, R\rangle]$. For more on this treatment of cardinals and ordinals, see Potter (2004, chs. 9–12).

choice.3 Comparability and Hartogs' Lemma

That's the plus side. Here's the minus side. Without Choice, things get *messy.* To see why, here is a nice result due to Hartogs (1915): sth:choice:hartogs:sec

Lemma choice.3 (*in* ZF). *For any set A, there is an ordinal α such that $\alpha \not\preceq A$* *sth:choice:hartogs:HartogsLemma*

Proof. If $B \subseteq A$ and $R \subseteq B^2$, then $\langle B, R\rangle \subseteq V_{\text{rank}(A)+4}$ by **??**. So, using Separation, consider:

$$C = \{\langle B, R\rangle \in V_{\text{rank}(A)+5} : B \subseteq A \text{ and } \langle B, R\rangle \text{ is a well-ordering}\}$$

Using Replacement and **??**, form the set:

$$\alpha = \{\text{ord}(B, R) : \langle B, R\rangle \in C\}.$$

By **??**, α is an ordinal, since it is a transitive set of ordinals. After all, if $\gamma \in \beta \in \alpha$, then $\beta = \text{ord}(B, R)$ for some $B \subseteq R$, whereupon $\gamma = \text{ord}(B_b, R_b)$ for some $b \in B$ by **??**, so that $\gamma \in \alpha$.

For reductio, suppose there is an injection $f\colon \alpha \to A$. Then, where:

$$\begin{aligned} B &= \text{ran}(f) \\ R &= \{\langle f(\alpha), f(\beta)\rangle \in A \times A : \alpha \in \beta\}. \end{aligned}$$

Clearly $\alpha = \text{ord}(B, R)$ and $\langle B, R\rangle \in C$. So $\alpha \in \alpha$, which is a contradiction. $\square$

[1] Which may have parameters.

 choice rev: 666b46f (2020-02-13) by OLP /

This entails a deep result:

Theorem choice.4 (*in* ZF). *The following claims are equivalent:*

sth:choice:hartogs:equivwo
1. *The Axiom of Well-Ordering*

sth:choice:hartogs:equivcompare
2. *Either $A \preceq B$ or $B \preceq A$, for any sets A and B*

Proof. *(1) ⇒ (2).* Fix A and B. Invoking (1), there are well-orderings $\langle A, R\rangle$ and $\langle B, S\rangle$. Invoking **??**, let $f\colon \alpha \to \langle A, R\rangle$ and $g\colon \beta \to \langle B, S\rangle$ be isomorphisms. By Trichotomy, either $\alpha \in \beta$ or $\alpha = \beta$ or $\beta \in \alpha$. In the first two cases $\alpha \subseteq \beta$, so $g \circ f^{-1}\colon A \to B$ is an injection, and hence $A \preceq B$. Similarly if $\beta \in \alpha$ then $B \preceq A$.

(1) ⇒ (2). Fix A; by Lemma choice.3 there is some ordinal β such that $\beta \not\preceq A$. Invoking (2), we have $A \preceq \beta$. So there is some injection $f\colon A \to \beta$, and we can use this injection to well-order the elements of A, by defining an order $\{\langle a, b\rangle \in A \times A : f(a) \in f(b)\}$. □

As an immediate consequence: if Well-Ordering fails, then some sets are *literally incomparable* with regard to their size. So, if Well-Ordering fails, then transfinite cardinal arithmetic will be messy. For example, we will have to abandon the idea that if A and B are infinite then $A \sqcup B \approx A \times B \approx M$, where M is the larger of A and B (see **??**). The problem is simple: if we cannot *compare* the size of A and B, then it is nonsensical to ask which is larger.

choice.4 The Well-Ordering Problem

sth:choice:woproblem:sec

Evidently rather a lot hangs on whether we accept Well-Ordering. But the discussion of this principle has tended to focus on an equivalent principle, the Axiom of Choice. So we will now turn our attention to that (and prove the equivalence).

In 1883, Cantor expressed his support for the Axiom of Well-Ordering, calling it "a law of thought which appears to me to be fundamental, rich in its consequences, and particularly remarkable for its general validity" (cited in Potter 2004, p. 243). But Cantor ultimately became convinced that the "Axiom" was in need of proof. So did the mathematical community.

The problem was "solved" by Zermelo in 1904. To explain his solution, we need some definitions.

Definition choice.5. A function f is a *choice function* iff $f(x) \in x$ for all $x \in \mathrm{dom}(f)$. We say that f is a *choice function for A* iff f is a choice function with $\mathrm{dom}(f) = A \setminus \{\emptyset\}$.

Intuitively, for every (non-empty) set $x \in A$, a choice function for A *chooses* a particular element, $f(x)$, from x. The Axiom of Choice is then:

Axiom (Choice). Every set has a choice function.

Zermelo showed that Choice entails well-ordering, and vice versa:

Theorem choice.6 (in ZF). *Well-Ordering and Choice are equivalent.* sth:choice:woproblem:thmwochoice

Proof. Left-to-right. Let A be a set of sets. Then $\bigcup A$ exists by the Axiom of Union, and so by Well-Ordering there is some $<$ which well-orders $\bigcup A$. Now let $f(x) =$ the $<$-least member of x. This is a choice function for A.

Right-to-left. Fix A. By Choice, there is a choice function, f, for $\wp(A) \setminus \{\emptyset\}$. Using Transfinite Recursion, define:

$$g(0) = f(A)$$
$$g(\alpha) = \begin{cases} \text{stop!} & \text{if } A = g[\alpha] \\ f(A \setminus g[\alpha]) & \text{otherwise} \end{cases}$$

(The indication to "stop!" is just a shorthand for what would otherwise be a more long-winded definition. That is, when $A = g[\alpha]$ for the first time, let $g(\delta) = A$ for all $\delta \leq \alpha$.) Note that we do stop, since otherwise we would have that $\alpha \prec \wp(A) \setminus \{\emptyset\}$ for every ordinal α, contradicting Lemma choice.3.

We do not stop until we have exhausted A. Since we stop, $\operatorname{ran}(g) = A$.

Since f is a choice function, for each α we have $g(\alpha) = f(A \setminus g[\alpha]) \in A \setminus g[\alpha]$; i.e., $g(\alpha) \notin g[\alpha]$. So if $g(\alpha) = g(\beta)$ then $g(\beta) \notin g[\alpha]$, i.e., $\beta \notin \alpha$, and similarly $\alpha \notin \beta$. So $\alpha = \beta$, by Trichotomy. So g is an injection.

Assembling these two facts, g is a bijection from some ordinal to A. Now g can be used to well-order A. □

So Well-Ordering and Choice stand or fall together. But the question remains: do they stand or fall?

choice.5 Countable Choice

It is easy to prove, without any use of Choice/Well-Ordering, that: sth:choice:countablechoice:sec

Lemma choice.7 (in $\mathbf{Z}^-$). *Every finite set has a choice function.*

Proof. Let $a = \{b_1, \ldots, b_n\}$. Suppose for simplicity that each $b_i \neq \emptyset$. So there are objects $c_1, \ldots, c_n$ such that $c_1 \in b_1, \ldots, c_n \in b_n$. Now by **??**, the set $\{\langle b_1, c_1\rangle, \ldots, \langle b_n, c_n\rangle\}$ exists; and this is a choice function for a. □

But matters get murkier as soon as we consider infinite sets. For example, consider this "minimal" extension to the above:

Countable Choice. Every *countable* set has a choice function.

This is a special case of Choice. And it transpires that this principle was invoked fairly frequently, without an obvious awareness of its use. Here are two nice examples.[2]

[2] Due to Potter (2004, §9.4) and Luca Incurvati.

Example 1. Here is a natural thought: for any set A, either $\omega \preceq A$, or $A \approx n$ for some $n \in \omega$. This is one way to state the intuitive idea, that every set is either finite or infinite. Cantor, and many other mathematicians, made this claim without proving it. Cautious as we are, we proved this in **??**. But in that proof we were working in **ZFC**, since we were assuming that any set A can be well-ordered, and hence that $|A|$ is guaranteed to exist. That is: we explicitly assumed Choice.

In fact, Dedekind (1888) offered his own proof of this claim, as follows:

Theorem choice.8 (*in* $\mathbf{Z}^-$ + *Countable Choice*). *For any A, either $\omega \preceq A$ or $A \approx n$ for some $n \in \omega$.*

Proof. Suppose $A \not\approx n$ for all $n \in \omega$. Then in particular for each $n < \omega$ there is subset $A_n \subseteq A$ with exactly 2^n elements. Using this sequence $A_0, A_1, A_2, \ldots$, we define for each n:

$$B_n = A_n \setminus (A_0 \cup A_1 \cup \ldots \cup A_{n-1}).$$

Now note the following

$$\begin{aligned} |A_0 \cup A_1 \cup \ldots \cup A_{n-1}| &\leq |A_0| + |A_1| + \ldots + |A_{n-1}| \\ &= 1 + 2 + \ldots + 2^{n-1} \\ &= 2^n - 1 \\ &< 2^n = |A_n| \end{aligned}$$

Hence each B_n has at least one member, c_n. Moreover, the B_ns are pairwise disjoint; so if $c_n = c_m$ then $n = m$. But every $c_n \in A$. So the function $f(n) = c_n$ is an injection $\omega \to A$. □

Dedekind did not flag that he had used Countable Choice. But, did *you* spot its use? Look again. (Really: Look again.)

The proof used Countable Choice twice. We used it once, to obtain our sequence of sets A_0, A_1, A_2, ... We then used it again to select our elements c_n from each B_n. Moreover, this use of Choice is ineliminable. Cohen (1966, p. 138) proved that the result fails if we have no version of Choice. That is: it is consistent with **ZF** that there are sets which are *incomparable* with ω.

Example 2. In 1878, Cantor stated that a countable union of countable sets is countable. He did not present a proof, perhaps indicating that he took the proof to be obvious. Now, cautious as we are, we proved a more general version of this result in **??**. But our proof explicitly assumed Choice. And even the proof of the less general result requires Countable Choice.

Theorem choice.9 (*in* $\mathbf{Z}^-$ + *Countable Choice*). *If A_n is countable for each $n \in \omega$, then $\bigcup_{n<\omega} A_n$ is countable.*

Proof. Without loss of generality, suppose that each $A_n \neq \emptyset$. So for each $n \in \omega$ there is a surjection $f_n \colon \omega \to A_n$. Define $f \colon \omega \times \omega \to \bigcup_{n<\omega} A_n$ by $f(m, n) = f_n(m)$. The result follows because $\omega \times \omega$ is countable (**??**) and f is a surjection. □

Did you spot the use of the Countable Choice? It is used to choose our sequence of functions f_0, f_1, f_2, ...[3] And again, the result fails in the absence of any Choice principle. Specifically, Feferman and Levy (1963) proved that it is consistent with **ZF** that a countable union of countable sets has cardinality $\beth_1$. But here is a much funnier statement of the point, from Russell:

> This is illustrated by the millionaire who bought a pair of socks whenever he bought a pair of boots, and never at any other time, and who had such a passion for buying both that at last he had $\aleph_0$ pairs of boots and $\aleph_0$ pairs of socks... Among boots we can distinguish right and left, and therefore we can make a selection of one out of each pair, namely, we can choose all the right boots or all the left boots; but with socks no such principle of selection suggests itself, and we cannot be sure, unless we assume the multiplicative axiom [i.e., in effect Choice], that there is any class consisting of one sock out of each pair. (Russell, 1919, p. 126)

In short, some form of Choice is needed to prove the following: If you have countably many pairs of socks, then you have (only) countably many socks. And in fact, without Countable Choice (or something equivalent), a countable union of countable sets can fail to be countable.

The moral is that Countable Choice was used repeatedly, without much awareness of its users. The philosophicaly question is: How could we *justify* Countable Choice?

An attempt at an intuitive justification might invoke an appeal to a supertask. Suppose we make the first choice in $1/2$ a minute, our second choice in $1/4$ a minute, ..., our n-th choice in $1/2^n$ a minute, ... Then within 1 minute, we will have made an ω-sequence of choices, and defined a choice function.

But what, really, could such a thought-experiment tell us? For a start, it relies upon taking this idea of "choosing" rather literally. For another, it seems to bind up mathematics in metaphysical possibility.

More important: it is not going to give us any justification for Choice *tout court*, rather than *mere* Countable Choice. For if we need *every* set to have a choice function, then we will need to be able to perform a "supertask of arbitrary ordinal length". Bluntly, that idea is laughable.

choice.6 Intrinsic Considerations about Choice

The broader question, then, is whether Well-Ordering, or Choice, or indeed the comparability of all sets as regards their size—it doesn't matter which—can be justified. sth:choice:justifications:sec

Here is an attempted *intrinsic* justification. Back in **??**, we introduced several principles about the hierarchy. One of these is worth restating:

[3] A similar use of Choice occurred in **??** when we "fixed a bijection f_β" for each $\beta \in \alpha$.

> *Stages-accumulate.* For any stage S, and for any sets which were formed *before* stage S: a set is formed at stage S whose members are exactly those sets. Nothing else is formed at stage S.

In fact, many authors have suggested that the Axiom of Choice can be justified via (something like) this principle. We will briefly provide a gloss on that approach.

We will start with a simple little result, which offers *yet another* equivalent for Choice:

sth:choice:justifications: choiceset

Theorem choice.10 (in ZF). *Choice is equivalent to the following principle. If the elements of A are disjoint and non-empty, then there is some C such that $C \cap x$ is a singleton for every $x \in A$. (We call such a C a choice set for A.)*

The proof of this result is straightforward, and we leave it as an exercise for the reader.

Problem choice.1. Prove Theorem choice.10. If you struggle, you can find a proof in (Potter, 2004, pp. 242–3).

The essential point is that a choice set for A is just the range of a choice function for A. So, to justify Choice, we can simply try to justify its equivalent formulation, in terms of the existence of choice sets. And we will now try to do exactly that.

Let A's elements be disjoint and non-empty. By *Stages-are-key* (see **??**), A is formed at some stage S. Note that all the elements of $\bigcup A$ are available before stage S. Now, by *Stages-accumulate*, for *any* sets which were formed before S, a set is formed whose members are exactly those sets. Otherwise put: every *possible* collections of earlier-available sets will exist at S. But it is certainly *possible* to select objects which could be formed into a choice set for A; that is just some very specific subset of $\bigcup A$. So: some such choice set exists, as required.

Well, that's a *very* quick attempt to offer a justification of Choice on intrinsic grounds. But, to pursue this idea further, you should read Potter's (2004, §14.8) neat development of it.

choice.7 The Banach-Tarski Paradox

sth:choice:banach: sec

We might also attempt to justify Choice, as Boolos attempted to justify Replacement, by appealing to *extrinsic* considerations (see **??**). After all, adopting Choice has many desirable consequences: the ability to compare every cardinal; the ability to well-order every set; the ability to treat cardinals as a particular kind of ordinal; etc.

Sometimes, however, it is claimed that Choice has *undesirable* consequences. Mostly, this is due to a result by Banach and Tarski (1924).

Theorem choice.11 (Banach-Tarski Paradox (in ZFC)). *Any ball can be decomposed into finitely many pieces, which can be reassembled (by rotation and transportation) to form two copies of that ball.*

At first glance, this is a bit amazing. Clearly the two balls have *twice* the volume of the original ball. But rigid motions—rotation and transportation—do not change volume. So it looks as if Banach-Tarski allows us to magick new matter into existence.

It gets worse.[4] Similar reasoning shows that a pea can be cut into finitely many pieces, which can then be reassembled (by rotation and transportation) to form an entity the shape and size of Big Ben.

None of this, however, holds in **ZF** on its own.[5] So we face a decision: reject Choice, or learn to live with the "paradox".

We're going to suggest that we should learn to live with the "paradox". Indeed, we don't think it's much of a paradox at all. In particular, we don't see why it is any more or less paradoxical than any of the following results:[6]

1. There are as many points in the interval $(0, 1)$ as in $\mathbb{R}$.
   Proof: consider $\tan(\pi(r - 1/2))$.

2. There are as many points in a line as in a square.
   See **??** and **??**.

3. There are space-filling curves.
   See **??** and **??**.

None of these three results require Choice. Indeed, we now just regard them as surprising, lovely, bits of mathematics. Maybe we should adopt the same attitude to the Banach-Tarski Paradox.

To be sure, a technical observation is required here; but it only requires keeping a level head. Rigid motions preserve volume. Consequently, the five[7] pieces into which the ball is decomposed cannot all be *measurable*. Roughly put, then, it makes no sense to assign a volume to these individual pieces. You should think of these as unpicturable, "infinite scatterings" of points. Now, maybe it is "weird" to conceive of such "infinitely scattered" sets. But their existence seems to fall out from the injunction, embodied in *Stages-accumulate*, that you should form *all possible* collections of earlier-available sets.

If none of that convinces, here is a final (extrinsic) argument in favour of embracing the Banach-Tarski Paradox. It immediately entails the best math joke of all time:

---

[4] See Tomkowicz and Wagon (2016, Theorem 3.12).

[5] Though Banach-Tarski can be proved with principles which are strictly weaker than Choice; see Tomkowicz and Wagon (2016, 303).

[6] Potter (2004, 276–7), Weston (2003, 16), Tomkowicz and Wagon (2016, 31, 308–9), make similar points, using other examples.

[7] We stated the Paradox in terms of "finitely many pieces". In fact, Robinson (1947) proved that the paradoxical decomposition can be achieved with *five* pieces (but no fewer). For a proof, see Tomkowicz and Wagon (2016, pp. 66–7).

Question. What's an anagram of "Banach-Tarski"?

Answer. "Banach-Tarski Banach-Tarski".

choice.8 Vitali's Paradox

sth:choice:vitali:sec

To get a real sense of whether the Banach-Tarski construction is acceptable or not, we should examine its *proof.* Unfortunately, that would require much more algebra than we can present here. However, we can offer some quick remarks which might shed some insight on the proof of Banach-Tarski,[8] by focussing on the following result:

sth:choice:vitali:vitaliparadox

Theorem choice.12 (Vitali's Paradox (in ZFC)). *Any circle can be decomposed into countably many pieces, which can be reassembled (by rotation and transportation) to form two copies of that circle.*

This is much easier to prove than the Banach–Tarski Paradox. We have called it "Vitali's Paradox", since it follows from Vitali's 1905 construction of an unmeasurable set. But the set-theoretic aspects of the proof of Vitali's Paradox and the Banach-Tarski Paradox are very similar. The essential difference between the results is just that Banach-Tarski considers a *finite* decomposition, whereas Vitali's Paradox onsiders a *countably infinite* decomposition. As Weston (2003) puts it, Vitali's Paradox "is certainly not nearly as striking as the Banach–Tarski paradox, but it does illustrate that geometric paradoxes can happen even in 'simple' situations."

Vitali's Paradox concerns a two-dimensional figure, a circle. So we will work on the plane, $\mathbb{R}^2$. Let R be the set of (clockwise) rotations of points around the origin by *rational* radian values between $[0, 2\pi)$. Here are some algebraic facts about R (if you don't understand the statement of the result, the proof will make its meaning clear):

sth:choice:vitali:rotationsgroupabelian

Lemma choice.13. *R forms an abelian group under composition of functions.*

Proof. Writing 0_R for the rotation by 0 radians, this is an identity element for R, since $\rho \circ 0_R = 0_R \circ \rho = \rho$ for any $\rho \in R$.

Every element has an inverse. Where $\rho \in R$ rotates by r radians, $\rho^{-1} \in R$ rotates by $2\pi - r$ radians, so that $\rho \circ \rho^{-1} = 0_R$.

Composition is associative: $(\tau \circ \sigma) \circ \rho = \tau \circ (\sigma \circ \rho)$ for any $\rho, \sigma, \tau \in R$

Composition is commutative: $\sigma \circ \rho = \rho \circ \sigma$ for any $\rho, \sigma \in R$. □

In fact, we can split our group R in half, and then use either half to recover the whole group:

sth:choice:vitali:disjointgroup

Lemma choice.14. *There is a partition of R into two disjoint sets, R_1 and R_2, both of which are a basis for R.*

[8] For a much fuller treatment, see Weston (2003) or Tomkowicz and Wagon (2016).

Proof. Let R_1 consist of the rotations by rational radian values in $[0, \pi)$; let $R_2 = R \setminus R_1$. By elementary algebra, $\{\rho \circ \rho : \rho \in R_1\} = R$. A similar result can be obtained for R_2. □

We will use this fact about groups to establish Theorem choice.12. Let $\mathbf{S}$ be the unit circle, i.e., the set of points exactly 1 unit away from the origin of the plane, i.e., $\{\langle r, s\rangle \in \mathbb{R}^2 : \sqrt{r^2 + s^2} = 1\}$. We will split $\mathbf{S}$ into parts by considering the following relation on $\mathbf{S}$:

$$r \sim s \textit{ iff } (\exists \rho \in R)\rho(r) = s.$$

That is, the points of $\mathbf{S}$ are linked by this relation iff you can get from one to the other by a rational-valued rotation about the origin. Unsurprisingly:

Lemma choice.15. *$\sim$ is an equivalence relation.*

Proof. Trivial, using Lemma choice.13. □

We now invoke Choice to obtain a set, C, containing exactly one member from each equivalence class of $\mathbf{S}$ under $\sim$. That is, we consider a choice function f on the set of equivalence classes,[9]

$$E = \{[r]_\sim : r \in \mathbf{S}\},$$

and let $C = \mathrm{ran}(f)$. For each rotation $\rho \in R$, the set $\rho[C]$ consists of the points obtained by applying the rotation ρ to each point in C. These next two results show that these sets cover the circle completely and without overlap:

Lemma choice.16. $\mathbf{S} = \bigcup_{\rho \in R} \rho[C]$. *sth:choice:vitali:vitalicover*

Proof. Fix $s \in \mathbf{S}$; there is some $r \in C$ such that $r \in [s]_\sim$, i.e., $r \sim s$, i.e., $\rho(r) = s$ for some $\rho \in R$. □

Lemma choice.17. *If $\rho_1 \neq \rho_2$ then $\rho_1[C] \cap \rho_2[C] = \emptyset$.* *sth:choice:vitali:vitalinooverlap*

Proof. Suppose $s \in \rho_1[C] \cap \rho_2[C]$. So $s = \rho_1(r_1) = \rho_2(r_2)$ for some $r_1, r_2 \in C$. Hence $\rho_2^{-1}(\rho_1(r_1)) = r_2$, and $\rho_2^{-1} \circ \rho_1 \in R$, so $r_1 \sim r_2$. So $r_1 = r_2$, as C selects exactly one member from each equivalence class under $\sim$. So $s = \rho_1(r_1) = \rho_2(r_1)$, and hence $\rho_1 = \rho_2$. □

We now apply our earlier algebraic facts to our circle:

Lemma choice.18. *There is a partition of $\mathbf{S}$ into two disjoint sets, D_1 and D_2, such that D_1 can be partitioned into countably many sets which can be rotated to form a copy of $\mathbf{S}$ (and similarly for D_2).* *sth:choice:vitali:pseudobanachtarski*

[9]Note: since R is enumerable, each element of E is enumerable. Since $\mathbf{S}$ is non-enumerable, it follows from Lemma choice.16 and **??** that E is non-enumerable. So this is a use of *un*countable Choice.

Proof. Using R_1 and R_2 from Lemma choice.14, let:

$$D_1 = \bigcup_{\rho \in R_1} \rho[C] \qquad\qquad D_2 = \bigcup_{\rho \in R_1} \rho[C]$$

This is a partition of $\mathbf{S}$, by Lemma choice.16, and D_1 and D_2 are disjoint by Lemma choice.17. By construction, D_1 can be partitioned into countably many sets, $\rho[C]$ for each $\rho \in R_1$. And these can be rotated to form a copy of $\mathbf{S}$, since $\mathbf{S} = \bigcup_{\rho \in R} \rho[C] = \bigcup_{\rho \in R_1} (\rho \circ \rho)[C]$ by Lemma choice.14 and Lemma choice.16. The same reasoning applies to D_2. □

This immediately entails Vitali's Paradox. For we can generate *two* copies of $\mathbf{S}$ from $\mathbf{S}$, just by splitting it up into countably many pieces (the various $\rho[C]$'s) and then rigidly moving them (simply rotate each piece of D_1, and first transport and then rotate each piece of D_2).

Let's recap the proof-strategy. We started with some algebraic facts about the group of rotations on the plane. We used this group to partition $\mathbf{S}$ into equivalence classes. We then arrived at a "paradox", by using Choice to select elements from each class.

We use exactly the same strategy to prove Banach–Tarski. The main difference is that the algebraic facts used to prove Banach–Tarski are significantly more complicated than those used to prove Vitali's Paradox. But those algebraic facts have nothing to do with Choice. We will summarise them quickly.

To prove Banach–Tarski, we start by establishing an analogue of Lemma choice.14: any *free group* can be split into four pieces, which intuitively we can "move around" to recover two copies of the whole group.[10] We then show that we can use two particular rotations around the origin of $\mathbb{R}^3$ to generate a free group of rotations, F.[11] (No Choice yet.) We now regard points on the surface of the sphere as "similar" iff one can be obtained from the other by a rotation in F. We then *use Choice* to select exactly one point from each equivalence class of "similar" points. Applying our division of F to the surface of the sphere, as in Lemma choice.18, we split that surface into four pieces, which we can "move around" to obtain two copies of the surface of the sphere. And this establishes (Hausdorff, 1914):

Theorem choice.19 (Hausdorff's Paradox (in ZFC)). *The surface of any sphere can be decomposed into finitely many pieces, which can be reassembled (by rotation and transportation) to form two disjoint copies of that sphere.*

A couple of further algebraic tricks are needed to obtain the full Banach-Tarski Theorem (which concerns not just the sphere's surface, but its interior too). Frankly, however, this is just icing on the algebraic cake. Hence Weston writes:

[10] The fact that we can use *four* pieces is due to Robinson (1947). For a recent proof, see Tomkowicz and Wagon (2016, Theorem 5.2). We follow Weston (2003, p. 3) in describing this as "moving" the pieces of the group.

[11] See Tomkowicz and Wagon (2016, Theorem 2.1).

> [...] the result on free groups is the *key step* in the proof of the Banach-Tarski paradox. From this point of view, the Banach-Tarski paradox is not a statement about $\mathbb{R}^3$ so much as it is a statement about the complexity of the group [of translations and rotations in $\mathbb{R}^3$]. (Weston, 2003, p. 16)

That is: whether we can offer a *finite* decomposition (as in Banach–Tarski) or a *countably infinite* decomposition (as in Vitali's Paradox) comes down to certain group-theoretic facts about working in two-dimension or three-dimensions.

Admittedly, this last observation slightly spoils the joke at the end of section choice.7. Since it is two dimensional, "Banach-Tarski" must be divided into a countable *infinity* of pieces, if one wants to rearrange those pieces to form "Banach-Tarski Banach-Tarski". To repair the joke, one must write in three dimensions. We leave this as an exercise for the reader.

One final comment. In section choice.7, we mentioned that the "pieces" of the sphere one obtains cannot be *measurable*, but must be unpicturable "infinite scatterings". The same is true of our use of Choice in obtaining Lemma choice.18. And this is all worth explaining.

Again, we must sketch some background (but this is *just* a sketch; you may want to consult a textbook entry on *measure*). To define a measure for a set X is to assign a value $\mu(E) \in \mathbb{R}$ for each E in some "σ-algebra" on X. Details here are not essential, except that the function μ must obey the principle of countable additivity: the measure of a countable union of disjoint sets is the sum of their individual measures, i.e., $\mu(\bigcup_{n<\omega} X_n) = \sum_{n<\omega} \mu(X_n)$ whenever the X_ns are disjoint. To say that a set is "unmeasurable" is to say that no measure can be suitably assigned. Now, using our R from before:

Corollary choice.20 (Vitali). *Let μ be a measure such that $\mu(\mathbf{S}) = 1$, and such that $\mu(X) = \mu(Y)$ if X and Y are congruent. Then $\rho[C]$ is unmeasurable for all $\rho \in R$.*

Proof. For reductio, suppose otherwise. So let $\mu(\sigma[C]) = r$ for some $\sigma \in R$ and some $r \in \mathbb{R}$. For any $\rho \in C$, $\rho[C]$ and $\sigma[C]$ are congruent, and hence $\mu(\rho[C]) = r$ for any $\rho \in C$. By Lemma choice.16 and Lemma choice.17, $\mathbf{S} = \bigcup_{\rho \in R} \rho[C]$ is a countable union of pairwise disjoint sets. So countable additivity dictates that $\mu(\mathbf{S}) = 1$ is the sum of the measures of each $\rho[C]$, i.e.,

$$1 = \mu(\mathbf{S}) = \sum_{\rho \in R} \mu(\rho[C]) = \sum_{\rho \in R} r$$

But if $r = 0$ then $\sum_{\rho \in R} r = 0$, and if $r > 0$ then $\sum_{\rho \in R} r = \infty$. □

Photo Credits

Bibliography

Banach, Stefan and Alfred Tarski. 1924. Sur la décomposition des ensembles de points en parties respectivement congruentes. *Fundamenta Mathematicae* 6: 244–77.

Cantor, Georg. 1878. Ein Beitrag zur Mannigfaltigkeitslehre. *Journal für die reine und angewandte Mathematik* 84: 242–58.

Cantor, Georg. 1883. *Grundlagen einer allgemeinen Mannigfaltigkeitslehre. Ein mathematisch-philosophischer Versuch in der Lehre des Unendlichen.* Leipzig: Teubner.

Cohen, Paul J. 1966. *Set Theory and the Continuum Hypothesis.* Reading, MA: Benjamin.

Dedekind, Richard. 1888. *Was sind und was sollen die Zahlen?* Braunschweig: Vieweg.

Ebbinghaus, Heinz-Dieter, Craig G. Fraser, and Akihiro Kanamori. 2010. *Ernst Zermelo. Collected Works*, vol. 1. Berlin: Springer-Verlag.

Feferman, Solomon and Azriel Levy. 1963. Independence results in set theory by Cohen's method II. *Notices of the American Mathematical Society* 10: 593.

Hartogs, Friedrich. 1915. Über das Problem der Wohlordnung. *Mathematische Annalen* 76: 438–43.

Hausdorff, Felix. 1914. Bemerkung über den Inhalt von Punktmengen. *Mathematische Annalen* 75: 428–34.

Potter, Michael. 2004. *Set Theory and its Philosophy.* Oxford: Oxford University Press.

Robinson, Raphael. 1947. On the decomposition of spheres. *Fundamenta Mathematicae* 34(1): 246–60.

Russell, Bertrand. 1919. *Introduction to Mathematical Philosophy.* London: Allen & Unwin.

Tomkowicz, Grzegorz and Stan Wagon. 2016. *The Banach-Tarski Paradox*. Cambridge: Cambridge University Press.

Vitali, Giuseppe. 1905. *Sul problema della misura dei gruppi di punti di una retta*. Bologna: Gamberini e Parmeggiani.

Weston, Tom. 2003. The Banach-Tarski paradox URL `http://people.math.umass.edu/~weston/oldpapers/banach.pdf`.

Zermelo, Ernst. 1904. Beweis, daß jede Menge wohlgeordnet werden kann. *Mathematische Annalen* 59: 514–516. English translation in (Ebbinghaus et al., 2010, pp. 115–119).