The Banach Tarski Theorem

Rich Schwartz

December 13, 2023

1 The Main Result

For all these notes, we consider subsets in Euclidean space, $\boldsymbol{R}^3$. Here is a special case of the Axiom of Choice, which we call the *Real Axiom of Choice*. Given a collection of disjoint subsets of $\boldsymbol{R}^3$ there exists a new set which has one point from each of our subsets. This seems innocent enough but it has very unsettling consequences.

Two subsets $A, B \subset \boldsymbol{R}^3$ are *isometric* if there is an isometry I of $\boldsymbol{R}^3$ such that $I(A) = B$. In this case we call B an *isometric copy* of A. Say that a subset $A^* \subset \boldsymbol{R}^3$ is *strange* if there are disjoint subsets $A_1^*, A_2^* \subset A^*$ which are each isometric to A^*. Let Δ_r denote the ball of radius r centered at the origin. Call A^* *substantial* if Δ_1 is contained in a finite union of isometric copies of A^*.

Below we deduce the following theorem from the Real Axiom of Choice.

Theorem 1.1 *Δ_1 contains a set which is both strange and substantial.*

Corollary 1.2 *Let $r > 0$. There is some $n = n(r)$ with the following properties: There are disjoint subsets $A_1, ..., A_n \subset \Delta_1$ and isometries $I_1, ..., I_n$ such that the union $\bigcup I_k(A_k)$ contains Δ_r.*

Proof: Let A^* be our strange and substantial subset of Δ_1. Since A^* contains 2 disjoint isometric copies of itself, it actually contiains 4 such, and 8 such, etc. Continuing this way, we see that A^* contains as many disjoint isometric topics of itself as we like. We first cover Δ_r with a finite number of isometric copies of Δ_1 and then we cover each of these isometric topics by isometric copies of A^* all taken from our supply contained in A^*. ♠

Corollary 1.3 *Let $r > 0$. There is some $n = n(r)$ with the following properties: Any $S \subset \Delta_r$ is the union of disjoint subsets $B_1, ..., B_n$ which have disjoint isometric copies $A_1, ..., A_n$ in Δ_1.*

Proof: Let $A_1, ..., A_n$ be as in Corollary 1.2. We let $B_k \subset S$ to be those points p such that $p \in I_k(A_k)$ but $p \notin I_j(A_j)$ for any $j < k$. Let $J_k = I_k^{-1}$. By construction $S = \bigcup B_k$ and $J_k(B_k) \subset A_k$. Hence the sets $J_k(B_k)$, for $k = 1, ..., n$, are disjoint subsets of Δ_1. ♠

Call a subset $S \subset \boldsymbol{R}^3$ *nice* if it contains a ball and is contained in a (bigger) ball. A *piecewise isometry* between sets S and T is a bijection $h : S \to T$ together with partitions $S = S_1 \cup ... \cup S_n$ and $T = T_1 \cup ... \cup T_n$ such that, for all i, we have $h(S_i) = T_i$ and the restriction of h to S_i is an isometry. If there is a piecewise isometry between sets, we can break one apart like a puzzle and reassemble it into the other.

Corollary 1.4 (Banach-Tarski) *Suppose that S, T are nice sets. Then there is a piecewise isometry from S to T.*

Proof: Using Corollary 1.3 and scaling we can find a piecewise isometry $f : S \to T' \subset T$, where T' is a subset of T. Likewise, we can find a piecewise isometry $g : T \to S' \subset S$, where S' is a subset of S. The rest of our proof is essentially the Schroeder-Bernstein Theorem.

We form a bipartite graph Γ. The white vertices of Γ are the points of S and the black vertices are the points of T. We draw an edge between each $p \in S$ and $f(p) \in T$, and an edge between each $q \in T$ and $g(q) \in S$. Each connected component γ of Γ is either a cycle of even length, a bi-infinite path, an infinite ray starting in S or an infinite ray starting in T. For all the cases except the last one, we use f to give a bijection between the white vertices of γ and the black ones. In the last case we use g^{-1} for this purpose. Call this bijection h_γ. The union of all these h_γ, taken over all components, gives a bijection $h : S \to T$. Not only that, h is a piecewise isometry because $h = f$ on one subset of S and $h = g^{-1}$ on the complementary subset. ♠

These results are pretty crazy because they seem to violate some principle of conservation of volume. One explanation is that the sets you use in the piecewise isometry are so complicated that they don't have a "volume". That is just the kind of crazy stuff you get if you accept the Real Axiom of Choice.

2 Theorem 1.1 modulo a detail

Let G denote the infinite group consisting of all words in the letters a, b, b^2, subject to the relations that a^2 and b^3 are the empty word. A typical element of G would be $abab^2 ababab^2$. The identity element is the empty word. The group law is concatenation. This group G is often denoted $\boldsymbol{Z}/2 * \boldsymbol{Z}/3$, and called the *free product of* $\boldsymbol{Z}/2$ *and* $\boldsymbol{Z}/3$. The group G is countable.

Let $SO(3)$ denote the group of rotations of $\boldsymbol{R}^3$. Here is the detail that I will take care of below.

Lemma 2.1 *There is an injective homomorphism* $\rho : G \to SO(3)$.

Each nontrivial $g \in G$ defines a line ℓ_g through the origin in $\boldsymbol{R}^3$, namely the axis of $\rho(g)$. Let Δ_1^* denote what we get by starting with Δ_1 and removing all these axes. So, Δ_1^* is the unit ball with countably many line segments through the origin removed.

We define a group action of G on Δ_1^*. The rule is that

$$g \cdot p = \rho(g)(p). \tag{1}$$

Lemma 2.2 *The stabilizer of each point is the trivial subgroup of* G.

Proof: Suppose that $g \cdot p = p$. Then p is fixed by the nontrivial rotation $\rho(g)$. But then p lies in the axis ℓ_g. But then $p \notin \Delta_1^*$. ♠

Thanks to this lemma and the Orbit Stabilizer Theorem, there is a bijection between G and any orbit. The **Real Axiom of Choice** lets us choose one element in each orbit. Thus, we can specify a particular bijection $G \leftrightarrow O$ between G and each orbit O.

We introduce 3 subsets of G:

- A consists of words starting with a, and the empty word.
- B_1 consists of words starting with b.
- B_2 consists of words starting with b^2.

This is a partition of G. For each orbit O we let O_A and O_{B_1} and O_{B_2} be the subsets of O corresponding to our partition under the bijection $G \leftrightarrow O$ we have chosen. Let A^* denote the union of O_A taken over all orbits. Likewise define B_1^* and B_2^*. We have a partition $\Delta_1^* = A^* \cup B_1^* \cup B_2^*$.

Lemma 2.3 *$b \cdot A^* = B_1^*$ and $b \cdot B_1^* = B_2^*$ and $b \cdot B_2^* = A^*$.*

Proof: It suffices to prove that this happens in each orbit. That is, we have to show for each orbit O that the action of b permutes O_A and O_{B_1} and O_{B_2}. We have chosen some $p \in O$ so that the bijection $G \leftrightarrow O$ is given by $g \to g \cdot p$. The set O_A consists of points of the form $g \cdot p$ when $p \in A$. By the group action property, $b \cdot g \cdot p = (bg) \cdot p$. Since $bg \in B_1$ we see that $b \cdot g \cdot p \in O_{B_1}$. This proves that $b \cdot O_A \subset O_{B_1}$. Now, a similar argument shows that $b \cdot O_{B_1} \subset O_{B_2}$ and $b \cdot O_{B_2} \subset O_A$. But then, since b^3 is the identity, we must have the stronger result that $b \cdot O_A = O_{B_1}$ and $b \cdot O_{B_1} = O_{B_2}$ and $b \cdot O_{B_2} = O_A$. ♠

Lemma 2.4 *$a \cdot B_k^* \subset A^*$.*

Proof: Same proof as the previous result. It suffices to prove this for each orbit. For each orbit, this boils down to the fact that left multiplication by a maps B_1 and B_2 both into A. ♠

The next two results finish the proof of Theorem 1.1.

Lemma 2.5 *A^* is strange.*

Proof: We let $A_k^* = a \cdot B_k^*$. The disjoint sets $A_1^*, A_2^* \subset A^*$ respectively are isometric to B_1^*, B_2^* because our group action is by isometries. Note that the three sets A^* and B_1^* and B_2^* are all isometric to each other, by Lemma 2.3. Hence A_1^* and A_2^* are each isometric to A^*. ♠

Lemma 2.6 *A^* is substantial.*

Proof: Since Δ_1^* is the union of 3 sets isometric to A^*, it suffices to prove that Δ_1^* is substantial. Since we only have countably many line segments to worry about, we can rotate Δ_1^* so that the missing segments of Δ_1^* are contained in the rotated copy and vice versa – except for the origin. Hence Δ_1 minus the origin is contained in the union of two isometric copies of Δ_1^*. Now we throw in a third (translated) one which contains the origin. ♠

3 Proof of Lemma 2.1

Strategy: Our proof is related to what physicists call *Wick rotation*. We consider objects indexed by a parameter $u \in \boldsymbol{C}$. When $u \in \boldsymbol{R}$, the objects correspond to homomorphisms $G \to SO(3)$ in disguise. When $u \in i\boldsymbol{R}$ is very near i, the corresponding object is another kind of homomorphism which we will easily see is injective. Then we use the miracle of polynomials to convert information about the imaginary case to information about the real case. We finish up using the Baire Category Theorem.

A *Moebius transformation* is a map of the form

$$g^*(z) = \frac{Az+B}{Cz+D}, \qquad AD - BC \neq 0. \tag{2}$$

These act on the *Riemann sphere* $\boldsymbol{C} \cup \infty$. They form a group which we call Γ. When $A + D = 0$ the map has order 2, and has 2 fixed points. We call such maps *involutions*.

Given any $u \in \boldsymbol{C} - \{0\}$ let I_u denote the involution whose fixed points are u and $-1/u$. Let R be the map given by $R(z) = \exp(2\pi i/3)z$. This map has order 3. Let $\rho_u^* : G \to \Gamma$ be the homomorphism such that $\rho_u^*(a) = I_u$ and $\rho_u^*(b) = R$. Here is the imaginary part of the story.

Lemma 3.1 (Ping Pong) *ρ_u^* is injective if $u \in i\boldsymbol{R}$ is very close to i.*

Proof: It is convenient to just write $a = I_u$ and $b = R$. When u is imaginary, a fixes points z, w with $w = 1/\overline{z}$. From this it follows that a commutes with the map $z \to 1/\overline{z}$, a map whose fixed point set is the unit circle S^1. But then a preserves the unit circle. So does b.

Let $\beta \subset S^1$ be the complement of the arc of length (say) 1/10 centered about 1. Let $\alpha \subset S^1$ be an intervel of length 1/100 centered at 1. By construction b and b^2 map $S^1 - \beta$ into β. If u is sufficiently close to 1 then a maps $S^1 - \alpha$ into α. Also $\alpha \subset S^1 - \beta$ and $\beta \subset S^2 - \alpha$. If we have any $g \in G$ then $\mu(g)$ maps any point in $S^1 - \alpha - \beta$ into $\alpha \cup \beta$ and hence is nontrivial. To see this in action, consider bab^2. Starting with $p \in S^1 - \alpha - \beta$ we get $b^2(p) \in \beta$ and $ab^2(p) \in \alpha$ and $bab^2(p) \in \beta$. ♠

Here is the real part of the story. Let S^2 denote the unit sphere in $\boldsymbol{R}^3$. *Stereographic projection* is a homeomorphism from S^2 to $\boldsymbol{C} \cup \infty$ given by

$$\Sigma(x_1, x_2, x_3) = \frac{x_1 + ix_2}{1 - x_3}. \tag{3}$$

Lemma 3.2 *When* $u \in \boldsymbol{R}$ *the homomorphism,* ρ_u^* *has the form* $\Sigma \circ \rho_u \circ \Sigma^{-1}$ *for some homomorphism* $\rho_u : G \to SO(3)$. *In particular,* ρ_u *is injective if and only if* ρ_u^* *is injective.*

Proof: Any element of $g \in SO(3)$ fixes antipodal points $\pm p$. Moreover, $z = \Sigma(p)$ and $w = \Sigma(-p)$ satisfy $z\overline{w} = -1$. This is exactly the relation satisfied by $z = u$ and $w = -1/u$ when $u \in \boldsymbol{R}$. These facts are the main reasons why the conjugate map $g^* = \Sigma \circ g \circ \Sigma^{-1} \in \Gamma$ is an involution fixing points z, w with $z\overline{w} = -1$. So, whenever $u \in \boldsymbol{R}$ the map I_u equals g^* for some $g \in SO(3)$. At the same time, the map R equals g^* for one of the two $g \in SO(3)$ which have order 3 and fix $\pm(0,0,1)$. In short, when $u \in \boldsymbol{R}$ the homomorphism ρ_u^* is conjugate to a homomorphism $\rho_u : G \to SO(3)$. ♠

To prove Lemma 2.1 it suffices to prove ρ_u^* is injective for some $u \in \boldsymbol{R}$. Now fix a nontrivial $g \in G$. Let $S_g \subset \boldsymbol{R}$ denote those u for which $\rho_u^*(g)$ is nontrivial. We just need to show that the total intersection $\bigcap S_g$ is nonempty.

Lemma 3.3 S_g *is open and dense in* $\boldsymbol{R}$.

Proof: The set S_g is open by continuity. Each matrix entry of $\rho_u^*(g)$ is a polynomial in u. Any polynomial that is constant on an open set of $\boldsymbol{R}$ is constant on $\boldsymbol{C}$. Hence, if S_g contains an open set then $\rho_u^*(g)$ is the identity for all $u \in \boldsymbol{C}$. But if we choose suitable $u \in i\boldsymbol{R}$ we get the injective homomorphisms from the Ping Pong Lemma. This is a contradiction. Hence $\boldsymbol{R} - S_g$ contains no open set. This shows that S_g is dense in $\boldsymbol{R}$. ♠

Lemma 3.4 (Baire Category) *The intersection of a countable collection of open dense subsets of* $\boldsymbol{R}$ *is dense in* $\boldsymbol{R}$.

Proof: Here is a sketch that is adapted to our situation. Choose any $p \in \boldsymbol{R}$. Since S_{g_1} is dense we can move p slightly so that $p \in S_{g_1}$. Since S_{g_2} is dense and S_{g_1} is open we can move p a tiny bit so that now it belongs to $S_{g_1} \cap S_{g_2}$. We move again so that p belongs too $S_{g_1} \cap S_{g_2} \cap S_{g_3}$. Etc. Taking care about the limit, we finally move to a point in the whole intersection. ♠

By the Baire Category Theorem, there is some nonzero $u \in \boldsymbol{R}$ contained in the intersection $\bigcap S_g$ taken over all nontrivial $g \in G$. But then ρ_u^* is injective. This completes the proof of Lemma 2.1.