8*. The Banach-Tarski paradox

May 3, 2012

The Banach-Tarski paradox is that a unit ball in Euclidean 3-space can be decomposed into finitely many parts which can then be reassembled to form two unit balls in Euclidean 3-space (maybe some parts are not used in these reassemblings). Reassembling is done using distance-preserving transformations. This is one of the most striking consequences of the axiom of choice, and is good background for the study of measure theory (of course the parts of the decomposition are not measurable). We give a proof of the theorem here without going into any side issues. We follow Wagon, **The Banach-Tarski paradox**, where variations and connections to measure theory are explained in full. The proof involves very little set theory, only the axiom of choice. Some third semester calculus and some linear algebra and simple group theory are used. Altogether the proof should be accessible to a first-year graduate student who has seen some applications of the axiom of choice.

We start with some preliminaries on geometry and linear algebra. The "reassembling" mentioned in the Banach-Tarski paradox is entirely done by rotations and translations. Given a line in 3-space and an angle ξ, we imagine the rotation about the given line through the angle ξ. Mainly we will be interested in rotations about lines that go through the origin. We indicate how to obtain the matrix representations of such rotations. First suppose that φ is the rotation about the z-axis counterclockwise through the angle ξ. Then, using polar coordinates,

$$\begin{aligned}
\varphi\begin{pmatrix} x \\ y \\ z \end{pmatrix} &= \varphi\begin{pmatrix} r\cos\theta \\ r\sin\theta \\ z \end{pmatrix} \\
&= \begin{pmatrix} r\cos(\theta+\xi) \\ r\sin(\theta+\xi) \\ z \end{pmatrix} \\
&= \begin{pmatrix} r\cos\theta\cos\xi - r\sin\theta\sin\xi \\ r\cos\theta\sin\xi + r\sin\theta\cos\xi \\ z \end{pmatrix} \\
&= \begin{pmatrix} x\cos\xi - y\sin\xi \\ x\sin\xi + y\cos\xi \\ z \end{pmatrix},
\end{aligned}$$

which gives the matrix representation of φ:

$$\begin{pmatrix} \cos\xi & -\sin\xi & 0 \\ \sin\xi & \cos\xi & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Similarly, the matrix representations of rotations counterclockwise through the angle ξ about the x- and y-axes are, respectively,

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\xi & -\sin\xi \\ 0 & \sin\xi & \cos\xi \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \cos\xi & 0 & \sin\xi \\ 0 & 1 & 0 \\ -\sin\xi & 0 & \cos\xi \end{pmatrix}.$$

Next, note that any rotation with respect to a line through the origin can be obtained as a composition of rotations about the three axes. This is easy to see using spherical coordinates. If l is a line through the origin and a point P different from the origin with spherical coordinates ρ, φ, θ, a rotation about l through an angle ξ can be obtained as follows: rotate about the z axis through the angle $-\theta$, then about the y-axis through the angle $-\varphi$ (thereby transforming l into the z-axis), then about the z-axis through the angle ξ, then back through φ about the y axis and through θ about the z-axis.

We want to connect this to linear algebra. Recall that a 3×3 matrix A is *orthogonal* provided that it is invertible and $A^T = A^{-1}$. Thus the matrices above are orthogonal. A matrix is orthogonal iff its columns form a basis for ${}^3\mathbb{R}$ consisting of mutually orthogonal unit vectors; this is easy to see. It is easy to check that a product of orthogonal matrices is orthogonal. Hence all of the rotations about lines through the origin are represented by orthogonal matrices.

Lemma 8.1. *If A is an orthogonal 3×3 real matrix and X and Y are 3×1 column vectors, then $(AX) \cdot (AY) = X \cdot Y$, where $\cdot$ is scalar multiplication.*

Proof. This is a simple computation:

$$(AX) \cdot (AY) = (AX)^T(AY) = X^T A^T AY = X^T A^{-1} AY = X^T Y = X \cdot Y. \qquad \square$$

It follows that any rotation about a line through the origin preserves distance, because $|P-Q| = \sqrt{(P-Q) \cdot (P-Q)}$ for any vectors P and Q. Such rotations have an additional property: their matrix representations have determinant 1. This is clear from the discussion above. It turns out that this additional property characterizes the rotations about lines through the origin (see M. Artin, **Algebra**), but we do not need to prove that. The following property of such matrices is very useful, however.

Lemma 8.2. *Suppose that A is an orthogonal 3×3 real matrix with determinant 1, A not the identity. Then there is a non-zero 3×1 matrix X such that for any 3×1 matrix Y,*

$$AY = Y \quad \text{iff} \quad \exists a \in \mathbb{R}[Y = aX].$$

Proof. Note that $A^T(A - I) = I - A^T = (I - A)^T$. Hence

$$-|A - I| = |I - A| = |(I - A)^T| = |A^T(A - I)| = |A^T|\,|A - I| = |A - I|.$$

It follows that $|A - I| = 0$. Hence the system of equations $(A - I)X = 0$ has a nontrivial solution, which gives the X we want. Namely, we then have $AX = X$, of course. Then $A(aX) = aAX = aX$. This proves $\Leftarrow$ in the equivalence of the lemma. It remains to do the converse. We may assume that X has length 1. Now we apply the Gram-Schmidt process to get a basis for ${}^3\mathbb{R}$ consisting of mutually orthogonal unit vectors, the first of which is X. We put them into a matrix B as column vectors, X the first column. Note that the first column of AB is X, since $AX = X$, and hence the first column of $B^{-1}AB$ is $(1\ 0\ 0)^T$. Since $B^{-1}AB$ is an orthogonal matrix, it follows because its columns are mutually orthogonal that it has the form

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & a & b \\ 0 & c & d \end{pmatrix}.$$

Now suppose that $AY = Y$. Let $B^{-1}Y = (u\ e\ f)^T$. Then

(1) $e = f = 0$.

For, suppose that (1) fails. Now $B^{-1}ABB^{-1}Y = B^{-1}AY = B^{-1}Y$, while a direct computation using the above form of $B^{-1}AB$ yields $B^{-1}ABB^{-1}Y = (u\quad ae + bf\quad ce + df)^T$. So we get the two equations

$$\begin{aligned} ae + bf &= e \\ ce + df &= f \end{aligned} \quad \text{or} \quad \begin{aligned} (a-1)e + bf &= 0 \\ ce + (d-1)f &= 0 \end{aligned}$$

Since (1) fails, it follows that the determinant $\begin{vmatrix} a-1 & b \\ c & d-1 \end{vmatrix}$ is 0. Thus $ad-a-d+1-bc = 0$. Now $B^{-1}AB$ has determinant 1, and its determinant is $ad-bc$, so we infer that $a+d = 2$. But $a^2 + c^2 = 1$ and $b^2 + d^2 = 1$ since the columns of $B^{-1}AB$ are unit vectors, so $|a| \leq 1$ and $|d| \leq 1$. Hence $a = d = 1$ and $b = c = 0$. So $B^{-1}AB$ is the identity matrix, so A is also, contradiction. Hence (1) holds after all.

From (1) we get $Y = B(u\ 0\ 0)^T = uX$, as desired. $\square$

One more remark on geometry: any rotation preserves distance. We already said this for rotations about lines through the origin. If l does not go through the origin, one can use a translation to transform it into a line through the origin, do the rotation then, and then translate back. Since translations clearly preserve distance, so arbitrary rotations preserve distance.

The first concrete step in the proof of the Banach, Tarski theorem is to describe a very special group of permutations of ${}^3\mathbb{R}$. Let φ be the counterclockwise rotation about the z-axis through the angle $\cos^{-1}(\frac{1}{3})$, and let ρ the similar rotation about the x-axis. The matrix representation of these rotations and their inverses is, by the above,

$$(1) \qquad \varphi^{\pm 1} = \begin{pmatrix} \frac{1}{3} & \mp\frac{2\sqrt{2}}{3} & 0 \\ \pm\frac{2\sqrt{2}}{3} & \frac{1}{3} & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \rho^{\pm 1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{3} & \mp\frac{2\sqrt{2}}{3} \\ 0 & \pm\frac{2\sqrt{2}}{3} & \frac{1}{3} \end{pmatrix}.$$

Let G_0 be the group of permutations of ${}^3\mathbb{R}$ generated by φ and ρ. By a *word* in φ and ρ we mean a finite sequence with elements in $\{\varphi, \varphi^{-1}, \rho, \rho^{-1}\}$. Given such a word $w = \langle \sigma_0, \ldots, \sigma_{m-1} \rangle$, we let $\overline{w}$ be the composition $\sigma_0 \circ \ldots \circ \sigma_{m-1}$. Further, we call w *reduced* if no two successive terms of w have any of the four forms $\langle \varphi, \varphi^{-1} \rangle$, $\langle \varphi^{-1}, \varphi \rangle$, $\langle \rho, \rho^{-1} \rangle$, or $\langle \rho^{-1}, \rho \rangle$.

Lemma 8.3. *If w is a reduced word of positive length, then $\overline{w}$ is not the identity.*

Proof. Suppose the contrary. Since $\rho \circ \overline{w} \circ \rho^{-1}$ is also the identity, we may assume that w ends with $\rho^{\pm 1}$ (on the right). [If w already ends with $\rho^{\pm 1}$, we do nothing. If it ends with $\varphi^{\pm 1}$, let $w' = \rho w \rho^{-1}$. Then w' is reduced, unless w has the form $\rho^{-1}w''$, in which case $w''\rho^{-1}$ is reduced, and still $\overline{w''\rho^{-1}} = \overline{w}$ =the identity.]

Since obviously $\rho^{\pm 1}$ is not the identity, w must have length at least 2. Now we claim

(1) For every terminal segment w' of w of nonzero even length the vector $\overline{w'}\,(1 \quad 0 \quad 0)^T$ has the form $(1/3^k)\,(a \quad b\sqrt{2} \quad c)^T$, with a divisible by 3 and b not divisible by 3.

We prove this by induction on the length of w'. First note that, by computation,

$$\rho\varphi = \frac{1}{9}\begin{pmatrix} 3 & -6\sqrt{2} & 0 \\ 2\sqrt{2} & 1 & -6\sqrt{2} \\ 8 & 2\sqrt{2} & 3 \end{pmatrix}; \quad \rho\varphi^{-1} = \frac{1}{9}\begin{pmatrix} 3 & 6\sqrt{2} & 0 \\ -2\sqrt{2} & 1 & -6\sqrt{2} \\ -8 & 2\sqrt{2} & 3 \end{pmatrix};$$

$$\rho^{-1}\varphi = \frac{1}{9}\begin{pmatrix} 3 & -6\sqrt{2} & 0 \\ 2\sqrt{2} & 1 & 6\sqrt{2} \\ -8 & -2\sqrt{2} & 3 \end{pmatrix}; \quad \rho^{-1}\varphi^{-1} = \frac{1}{9}\begin{pmatrix} 3 & 6\sqrt{2} & 0 \\ -2\sqrt{2} & 1 & 6\sqrt{2} \\ 8 & -2\sqrt{2} & 3 \end{pmatrix}.$$

Now we proceed by induction. For w' of length 2 we have

$$\rho\varphi\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \frac{1}{9}\begin{pmatrix} 3 \\ 2\sqrt{2} \\ 8 \end{pmatrix}; \quad \rho\varphi^{-1}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \frac{1}{9}\begin{pmatrix} 3 \\ -2\sqrt{2} \\ -8 \end{pmatrix};$$

$$\rho^{-1}\varphi\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \frac{1}{9}\begin{pmatrix} 3 \\ 2\sqrt{2} \\ -8 \end{pmatrix}; \quad \rho^{-1}\varphi^{-1}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \frac{1}{9}\begin{pmatrix} 3 \\ -2\sqrt{2} \\ 8 \end{pmatrix};$$

hence (1) holds in this case. The induction step:

$$\rho\varphi\left(\frac{1}{3^k}\begin{pmatrix} a \\ b\sqrt{2} \\ c \end{pmatrix}\right) = \frac{1}{3^{k+2}}\begin{pmatrix} 3a-12b \\ 2\sqrt{2}a + b\sqrt{2} - 6\sqrt{2}c \\ 8a+4b+3c \end{pmatrix};$$

$$\rho\varphi^{-1}\left(\frac{1}{3^k}\begin{pmatrix} a \\ b\sqrt{2} \\ c \end{pmatrix}\right) = \frac{1}{3^{k+2}}\begin{pmatrix} 3a+12b \\ -2\sqrt{2}a + b\sqrt{2} - 6\sqrt{2}c \\ -8a+4b+3c \end{pmatrix};$$

$$\rho^{-1}\varphi\left(\frac{1}{3^k}\begin{pmatrix} a \\ b\sqrt{2} \\ c \end{pmatrix}\right) = \frac{1}{3^{k+2}}\begin{pmatrix} 3a-12b \\ 2\sqrt{2}a + b\sqrt{2} + 6\sqrt{2}c \\ -8a-4b+3c \end{pmatrix};$$

$$\rho^{-1}\varphi^{-1}\frac{1}{3^{k+2}}\left(\frac{1}{3^k}\begin{pmatrix} a \\ b\sqrt{2} \\ c \end{pmatrix}\right) = \begin{pmatrix} 3a+12b \\ -2\sqrt{2}a + b\sqrt{2} + 6\sqrt{2}c \\ 8a-4b+3c \end{pmatrix}.$$

So, our assertion (1) is true. If w itself is of even length, then a contradiction has been reached, since b is not divisible by 3. If w is of odd length, then the following shows that the second entry of $\overline{w}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}^T$ is nonzero, still a contradiction:

$$\varphi\left(\frac{1}{3^k}\begin{pmatrix} a \\ b\sqrt{2} \\ c \end{pmatrix}\right) = \frac{1}{3^{k+1}}\begin{pmatrix} a-4b \\ 2\sqrt{2}a + b\sqrt{2} \\ 3c \end{pmatrix};$$

$$\varphi^{-1}\left(\frac{1}{3^k}\begin{pmatrix} a \\ b\sqrt{2} \\ c \end{pmatrix}\right) = \frac{1}{3^{k+1}}\begin{pmatrix} a+4b \\ -2\sqrt{2}a+b\sqrt{2} \\ 3c \end{pmatrix}.$$

This finishes the proof of Lemma 8.3 $\square$

This lemma really says that G_0 is (isomorphic to) the free group on two generators. But we do not need to go into that. We do need the following corollary, though.

Corollary 8.4. *For every $g \in G_0$ there is a unique reduced word w such that $g = \overline{w}$.*

Proof. Suppose that w and w' both work, and $w \neq w'$. Say $w = \langle \sigma_0, \ldots, \sigma_{m-1} \rangle$ and $w' = \langle \tau_0, \ldots, \tau_{n-1} \rangle$. If one is a proper segment of the other, say by symmetry w is a proper segment of w', then

$$\begin{aligned} g &= \overline{w} = \sigma_0 \circ \ldots \circ \sigma_{m-1} \\ &= \overline{w'} = \tau_0 \circ \ldots \circ \tau_{n-1}; \end{aligned}$$

since $\sigma_i = \tau_i$ for all $i < m$, we obtain $I = \tau_m \circ \ldots \circ \tau_{n-1}$, I the identity. But $\langle \tau_m, \ldots, \tau_{n-1} \rangle$ is reduced, contradicting 8.3.

Thus neither w nor w' is a proper initial segment of the other. Hence there is an $i < \min(m, n)$ such that $\sigma_i \neq \tau_i$, while $\sigma_j = \tau_j$ for all $j < i$ (maybe $i = 0$). But then we have by cancellation $\sigma_i \circ \ldots \circ \sigma_{m-1} = \tau_i \circ \ldots \circ \tau_{n-1}$, so $\tau_{n-1}^{-1} \circ \ldots \circ \tau_i^{-1} \sigma_i \circ \ldots \circ \sigma_{m-1} = I$. But since $\sigma_i \neq \tau_i$, the word $\langle \tau_{n-1}^{-1}, \ldots, \tau_i^{-1}, \sigma_i, \ldots, \sigma_{m-1} \rangle$ is reduced, again contradicting 8.3. $\square$

If G is a group and X is a set, we say that G *acts on* X if there is a homomorphism from G into the group of all permutations of X. Usually this homomorphism will be denoted by $\check{}$, so that $\check{g}$ is the permutation of X corresponding to $g \in G$. (Most mathematicians don't even use $\check{}$, using the same symbol for elements of the group and for the image under the homomorphism.) An important example is: any group G acts on itself by left multiplication. Thus for any $g \in G$, $\check{g} : G \to G$ is defined by $\check{g}(h) = g \cdot h$, for all $h \in G$.

Let G act on a set X, and let $E \subseteq X$. Then we say that E is *G-paradoxical* if there are positive integers m, n and pairwise disjoint subsets $A_0, \ldots, A_{m-1}, B_0, \ldots, B_{n-1}$ of E, and elements $\langle g_i : i < m \rangle$ and $\langle h_i : i < n \rangle$ of G such that $E = \bigcup_{i<m} \check{g}_i[A_i]$ and $E = \bigcup_{j<n} \check{h}_j[B_j]$. Note that this comes close to the Banach-Tarski formulation, except that the sets X and E are unspecified.

Lemma 8.5. *G_0, acting on itself by left multiplication, is G_0-paradoxical.*

Proof. If σ is one of $\phi^{\pm 1}$, $\rho^{\pm 1}$, we denote by $W(\sigma)$ the set of all reduced words beginning on the left with σ, and $\overline{W}(\sigma) = \{\overline{w} : w \in W(\sigma)\}$. Thus, obviously,

$$G_0 = \{I\} \cup \overline{W}(\phi) \cup \overline{W}(\phi^{-1}) \cup \overline{W}(\rho) \cup \overline{W}(\rho^{-1}),$$

where I is the identity element of G_0. These five sets are pairwise disjoint by 8.4. Thus the lemma will be proved, with $m = n = 2$, by proving the following two statements:

(1) $G_0 = \overline{W}(\phi) \cup \check{\phi}[\overline{W}(\phi^{-1})]$.

To see this, suppose that $g \in G_0$ and $g \notin \overline{W}(\phi)$. Write $g = \overline{w}$, w a reduced word. Then w does not start with ϕ. Hence $\phi^{-1}w$ is still a reduced word, and $g = \phi \circ \phi^{-1} \circ \overline{w} \in \check{\phi}[\overline{W}(\phi^{-1})]$, as desired.

(2) $G_0 = \overline{W}(\rho) \cup \check{\rho}[\overline{W}(\rho^{-1})]$.

The proof is just like for (1). □

We need two more definitions, given that G acts on a set X. For each $x \in X$, the *G-orbit* of x is $\{\check{g}(x) : g \in G\}$. The set of G-orbits forms a partition of X. We say that G *is without non-trivial fixed points* if for every non-identity $g \in G$ and every $x \in X$, $\check{g}(x) \neq x$.

The following lemma is the place in the proof of the Banach-Tarski paradox where the axiom of choice is used. Don't jump to the conclusion that the proof is almost over; our group G_0 above has non-trivial fixed points, and so does not satisfy the hypothesis of the lemma. Some trickery remains to be done even after this lemma. [For example, all points on the z-axis are fixed by φ.]

Lemma 8.6. *Suppose that G is G-paradoxical and acts on a set X without non-trivial fixed points. Then X is G-paradoxical.*

Proof. Let

$$A_0, \ldots, A_{m-1}, B_0, \ldots, B_{n-1}, g_0, \ldots, g_{m-1}, h_0, \ldots, h_{n-1}$$

be as in the defintion of paradoxical. By AC, let M be a subset of X having exactly one element in common with each G-orbit. Then we claim:

(1) $\langle \check{g}[M] : g \in G \rangle$ is a partition of X.

First of all, obviously each set $\check{g}[M]$ is nonempty. Next, their union is X, since for any $x \in X$ there is a $y \in M$ which is in the same G-orbit as x, and this yields a $g \in G$ such that $x = \check{g}(y)$ and hence $x \in \check{g}[M]$. Finally, if g and h are distinct elements of G, then $\check{g}[M]$ and $\check{h}[M]$ are disjoint. In fact, otherwise let y be a common element. Say $\check{g}(x) = y$, $x \in M$, and $\check{h}(z) = y$, $z \in M$. Then clearly x and z are in the same G-orbit, so $x = z$ since they are "both" in M. Then $(g^{-1} \cdot h)\check{}(z) = z$ and $g^{-1} \cdot h$ is not the identity, contradicting the no non-trivial fixed point assumption. So, (1) holds.

Now let $A_i^* = \bigcup_{g \in A_i} \check{g}[M]$ and $B_j^* = \bigcup_{g \in B_j} \check{g}[M]$, for all $i < m$ and $j < n$.

(2) $A_i^* \cap A_k^* = 0$ if $i < k < m$.

In fact, suppose that $x \in A_i^* \cap A_k^*$. Then we can choose $g \in A_i$ and $h \in A_k$ such that $x \in \check{g}[M] \cap \check{h}[M]$. But $g \neq h$ since $A_i \cap A_k = 0$, so this contradicts (1). Similarly the following two conditios hold:

(3) $B_i^* \cap B_k^* = 0$ if $i < k < n$.
(4) $A_i^* \cap B_j^* = 0$ if $i < m$ and $j < n$.

(5) $\bigcup_{i<m} \check{g}_i[A_i^*] = X$.

For, let $x \in X$. Say by (1) that $x \in \check{g}[M]$. Then by the choice of the A_i's there is an $i < m$ such that $g \in \check{g}_i[A_i]$. Say $h \in A_i$ and $g = \check{g}_i(h) = g_i \cdot h$. Since $x \in \check{g}[M]$, say $x = \check{g}(m)$

with $m \in M$. Then $x = (g_i \cdot h)\check{}(m) = \check{g}_i(\check{h}(m))$. Now $\check{h}(m) \in \check{h}[M] \subseteq A_i^*$, so $x \in \check{g}_i[A_i^*]$, as desired in (5).

(6) $\bigcup_{i<n} \check{h}_i[B_i^*] = X$.

This is proved similarly. $\square$

Let $S^2 = \{x \in {}^3\mathbb{R} : |x| = 1\}$ be the usual unit sphere. Now we can prove the first paradoxical result leading to the Banach-Tarski paradox:

Theorem 8.7. (Hausdorff) *There is a countable $D \subseteq S^2$ such that $S^2 \backslash D$ is G_0-paradoxical.*

Proof. Let D be the set of all fixed points of non-identity elements of G_0. By 8.2, D is countable. Now we claim that if $\sigma \in G_0$, then $\sigma[S^2\backslash D] = S^2\backslash D$. For, assume that $x \in S^2\backslash D$ and $\sigma(x) \in D$. Say $\tau \in G_0$, τ not the identity, and $\tau(\sigma(x)) = \sigma(x)$. Then $\sigma^{-1}\tau\sigma(x) = x$. Now $\sigma^{-1} \circ \tau \circ \sigma$ is not the identity, since τ isn't, so $x \in D$, contradiction. This proves that $\sigma[S^2\backslash D] \subseteq S^2\backslash D$. This holds for any $\sigma \in G_0$, in particular for σ^{-1}, and applying σ to that inclusion yields $S^2\backslash D \subseteq \sigma[S^2\backslash D]$, so the desired equality holds.

Thus G_0 acts on $S^2\backslash D$ without non-trivial fixed points. So by 8.5 and 8.6, $S^2\backslash D$ is G_0-paradoxical. $\square$

Let us see how far we have to go now. This theorem only looks at the sphere, not the ball. A countable subset is ignored. Since the sphere is uncountable, this makes the result close to what we want. But actually there is a countable subset of the sphere which is dense on it. [Take points whose spherical coordinates are rational.]

For the next step we need a new notion. Suppose that G is a group acting on a set X, and $A, B \subseteq X$. We say that *A and B are finitely G-equidecomposable* if A and B can be decomposed into the same number of parts, each part of A being carried into the corresponding part of B by an element of G. In symbols, there is a positive integer n such that there are partitions $A = \bigcup_{i<n} A_i$ and $B = \bigcup_{i<n} B_i$ and members $g_i \in G$ for $i < n$ such that $\check{g}_i[A_i] = B_i$ for all $i < n$. We then write $A \sim_G B$.

Lemma 8.8. *If G acts on a set X, then $\sim_G$ is an equivalence relation on $\mathscr{P}(X)$.*

Proof. Obviously $\sim_G$ is reflexive on $\mathscr{P}(X)$ and is symmetric. Now suppose that $A \sim_G B \sim_G C$. Then we get partitions $A = \bigcup_{i<m} A_i$ and $B = \bigcup_{i<m} B_i$ with elements $g_i \in G$ such that $\check{g}_i[A_i] = B_i$ for all $i < m$; and partitions $B = \bigcup_{j<n} B'_j$ and $C = \bigcup_{j<n} C_j$ with elements $h_i \in G$ such that $\check{h}_j[B'_j] = C_j$ for all $j < n$. Now for all $i < m$ and $j < n$ let $B_{ij} = B_i \cap B'_j$, $A_{ij} = \check{g}_i{}^{-1}[B_{ij}]$, and $C_{ij} = \check{h}_j[B_{ij}]$. Then $A = \bigcup_{i<m,j<n} A_{ij}$ is a partition of A, $C = \bigcup_{i<m,j<n} C_{ij}$ is a partition of C, and $(h_j \cdot g_i)\check{}[A_{ij}] = C_{ij}$. Some of the B_{ij} may be empty; eliminating the empty ones yields the desired nonemptiness of members of the partitions. $\square$

Lemma 8.9. *Suppose that G acts on X, E and E' are finitely G-equidecomposable subsets of X, and E is G-paradoxical. Then also E' is G-paradoxical.*

Proof. Because E is G-paradoxical, we can find pairwise disjoint subsets

$$A_0, \ldots, A_{m-1}, B_0, \ldots, B_{n-1}$$

of E and corresponding elements $g_0, \ldots, g_{m-1}, h_0, \ldots, h_{n-1}$ of G such that

$$E = \bigcup_{i<m} \check{g}_i[A_i] = \bigcup_{j<n} \check{h}_j[B_i].$$

And because E and E' are finitely G=equidecomposable we can find partitions $E = \bigcup_{k<p} C_k$ and $E' = \bigcup_{k<p} D_k$ with elements $f_i \in G$ such that $\check{f}_i[C_k] = D_k$ for all $k < p$. Then the following sets are pairwise disjoint: $A_i \cap \check{g_i}^{-1}[C_k]$ for $i < m$ and $k < p$, and $B_j \cap \check{h_j}^{-1}[C_k]$ for $j < n$ and $k < p$. And

$$\begin{aligned}
E' &= \bigcup_{k<p} D_k = \bigcup_{k<p} \check{f}_k[C_k] \\
&= \bigcup_{k<p} \check{f}_k[C_k \cap \bigcup_{i<m} \check{g}_i[A_i]] \\
&= \bigcup_{k<p,i<m} \check{f}_k[C_k \cap \check{g}_i[A_i]] \\
&= \bigcup_{k<p,i<m} (f_k \cdot g_i)\check{\ }[A_i \cap \check{g_i}^{-1}[C_k]],
\end{aligned}$$

and similarly

$$E' = \bigcup_{k<p,j<n} (f_k \cdot h_j)\check{\ }[B_i \cap \check{h_j}^{-1}[C_k]]. \qquad \square$$

Lemma 8.10. *Let D be a countable subset of S^2. Then there is a rotation σ with respect to a line through the origin such that if G_1 is the group of transformations of ${}^3\mathbb{R}$ generated by σ, then S^2 and $S^2\backslash D$ are G_1-equidecomposable.*

Proof. For each $d \in D$ let $f(d)$ be the line through the origin and d. Then f maps D into the set L of all lines through the origin, and the range of f is countable. But L itself is uncountable: for example, for each $\theta \in [0, \pi]$ one can take the line through the origin and $(\cos\theta, \sin\theta, 0)$. Hence there is a line $l \in L$ not in the range of f. This means that l does not pass through any point of D. Fix a direction in which to take rotations about l.

Note that if P and Q are distinct points of D, then there is at most one rotation about l which takes P to Q and is between 0 and 2π; this will be denoted by ψ_{PQ}, if it exists. Now let A consist of all $\theta \in (0, 2\pi)$ such that there is a positive integer n and a $P \in D$ such that if σ is the rotation about l through the angle $n\theta$, then $\sigma(P) \in D$. We claim that A is countable. For, if $P, Q \in D$, ψ_{PQ} is defined, $n \in \omega\backslash\{0\}$, $k \in \omega$, and $0 < \frac{1}{n}(\psi_{PQ} + 2\pi k) < 2\pi$, then $\frac{1}{n}(\psi_{PQ} + 2\pi k) \in A$; and every member of A can be obtained this way. [Given $\theta \in A$, we have $n\theta = \psi_{PQ} + 2\pi k$ for some $P, Q \in D$ and $n, k \in \omega$.] This really defines a function from $D \times D \times (\omega\backslash\{0\}) \times \omega$ onto A, so A is, indeed, countable. We choose $\theta \in (0, 2\pi)\backslash A$, and take the rotation σ about l through the angle θ. Let $\overline{D} = \bigcup_{n\in\omega} \sigma^n[D]$. The choice of σ says that $\sigma^n[D] \cap D = 0$ for every positive integer n. Hence if $n < m < \omega$ we have $\sigma^n[D] \cap \sigma^m[D] = 0$, since

$$\sigma^n[D] \cap \sigma^m[D] = \sigma^n[D \cap \sigma^{m-n}[D]] = \sigma^n[0] = 0.$$

Note that $\sigma[\overline{D}] = \overline{D} \backslash D$. Hence

$$S^2 = \overline{D} \cup (S^2 \backslash \overline{D}) \sim_{G_1} \sigma[\overline{D}] \cup (S^2 \backslash \overline{D}) = S^2 \backslash D. \qquad \square$$

Now let G_2 be the group of permutations of ${}^3\mathbb{R}$ generated by $\{\varphi, \rho, \sigma\}$. We now have the first form of the Banach-Tarski paradox:

Theorem 8.11. (Banach, Tarski) *S^2 is G_2-paradoxical.* $\square$

One can loosely state this theorem as follows: one can decompose S^2 into a finite number of pieces, rotate some of these pieces finitely many times with respect to certain lines through the origin to reassemble S^2, and then similarly transform some of the remaining pieces to also reassemble S^2. The rotations are of three kinds: the very specific rotations φ and ρ defined at the beginning of this section, and the rotation σ in the preceding proof, for which we do not have an explicit description. One can apply the inverses of these rotations as well. After doing the second reassembling, one can apply a translation to make that copy of S^2 disjoint from the first copy.

Finally, let $B = \{x \in {}^3\mathbb{R} : |x| \leq 1\}$ be the unit ball in 3-space. Let G_3 be the group generated by φ, ρ, σ, and the rotation τ about the line determined by $(0, 0, \frac{1}{2})$ and $(1, 0, \frac{1}{2})$, through the angle $\pi/\sqrt{2}$. Note that by the considerations at the beginning of this section, τ consists of the translation $\begin{pmatrix} x & y & z \end{pmatrix}^T \mapsto \begin{pmatrix} x & y & z - \frac{1}{2} \end{pmatrix}^T$, followed by the rotation through $\frac{\pi}{\sqrt{2}}$ about the x-axis, followed by the translation $\begin{pmatrix} x & y & z \end{pmatrix}^T \mapsto \begin{pmatrix} x & y & z + \frac{1}{2} \end{pmatrix}^T$.

Lemma 8.12. *For any positive integer k,*

$$\tau^k \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{1}{2} \sin\left(\frac{k\pi}{\sqrt{2}}\right) \\ -\frac{1}{2} \cos\left(\frac{k\pi}{\sqrt{2}}\right) + \frac{1}{2} \end{pmatrix}.$$

Hence $\tau^p \begin{pmatrix} 0 & 0 & 0 \end{pmatrix}^T \neq \begin{pmatrix} 0 & 0 & 0 \end{pmatrix}^T$ for every positive integer p.

Proof. The first equation is easily proved by induction on k. Then the second inequality follows since for any positive integer p, the argument $\frac{p\pi}{\sqrt{2}}$ is never equal to $m\pi$ for any integer m, since $\sqrt{2}$ is irrational. $\square$

Theorem 8.13. (Banach, Tarski) *B is G_3-paradoxical.*

Proof. By 8.11 there are pairwise disjoint subsets A_i and B_j of S^2 and members g_i, h_j of G_2 for $i < m$ and $j < n$ such that $S^2 = \bigcup_{i<m} g_i[A_i] = \bigcup_{j<n} h_j[B_j]$. For each $i < m$ and $j < n$ let $A'_i = \{\alpha x : x \in A_i,\ 0 < \alpha \leq 1\}$ and $B_j = \{\alpha x : x \in B_j,\ 0 < \alpha \leq 1\}$. Then

(1) The A'_i's and B'_j's are pairwise disjoint.

For example, suppose that $y \in A'_i \cap B'_j$. Then we can write $y = \alpha x$ with $x \in A_i$, $0 < \alpha \leq 1$, and also $y = \beta z$ with $z \in B_j$, and $0 < \beta \leq 1$. Hence $|y| = \alpha = \beta$. Hence $x = z$, contradiction.

(2) $B \backslash \{0\} = \bigcup_{i<m} g_i[A'_i] = \bigcup_{j<n} h_j[B'_j]$.

In fact, let $y \in B\backslash\{0\}$. Let $x = y/|y|$. Then $x \in S^2$, so there is an $i < m$ such that $x \in g_i[A_i]$. Say that $x = g_i(z)$ with $z \in A_i$. Then $|y|z \in A'_i$, and $g_i(|y|z) = |y|g_i(z) = |y|x = y$. So $y \in g_i[A'_i]$. This proves the first equality in (2), and the second equality is proved similarly.

So far, we have shown that $B\backslash\{0\}$ is G_2-paradoxical. Now we show that B and $B\backslash\{0\}$ are finitely G_3-equidecomposable, which will finish the proof. By lemma 8.12 we have

$$B = D \cup (B\backslash D) \sim_{G_3} \tau[D] \cup (B\backslash D) = B\backslash\{0\}.$$

This proves the desired equidecomposablity. □

As in the case of S^2, a translation can be made if one wants one copy of B to be disjoint from the other.