Chapter 11
Paradoxical Decompositions

SET-THEORETIC PARADOXES OF HAUSDORFF AND BANACH–TARSKI

Overview. In Chap. 10 we used the fixed-point theorem of Markov and Kakutani to show that every abelian group G is "amenable" in the sense that there is a G-invariant mean on the vector space $B(G)$ of bounded, real-valued functions on G. We observed that existence of such a "mean" is equivalent to existence of a finitely additive probability "measure" on $\mathscr{P}(G)$, the algebra of all subsets of G, and we asked if *every* group turns out to be amenable. We showed that the free group F_2 on two generators is *not* amenable by finding within F_2 four pairwise disjoint subsets that could be reassembled, using only group motions, into *two* copies of F_2.

Now we'll see how this "paradoxical" property of F_2, along with the Axiom of Choice, leads to astonishing results in set theory, most notably the famous Banach–Tarski Paradox, often popularly phrased as: *Each (three dimensional) ball can be partitioned into a finite collection of subsets which can then be reassembled, using only rigid motions, into* two *copies of itself.* Even more striking: given two bounded subsets of $\mathbb{R}^3$ with nonvoid interior, each can be partitioned into a finite collection of subsets that can be rigidly reassembled into the other. For this result the fixed-point theorem of Knaster and Tarski (Theorem 1.2) makes another appearance, this time to prove a far-reaching generalization of the Schröder–Bernstein Theorem.

Prerequisites. Elementary properties of: sets, groups, matrices.

11.1 Paradoxical Sets

To establish non-amenability for the free group F_2 on the two generators a and b (Theorem 10.12) we observed that its pairwise disjoint family of subsets $\mathscr{W} = \{W(a), W(a^{-1}), W(b), W(b^{-1})\}$ could be "F_2-reassembled" into *two* copies of F_2 in the sense that

$$F_2 = W(a) \uplus aW(a^{-1}) = W(b) \uplus bW(b^{-1}),$$

J.H. Shapiro, *A Fixed-Point Farrago*, Universitext,
DOI 10.1007/978-3-319-27978-7_11

where the symbol "$\uplus$" denotes "union of pairwise disjoint sets," and $W(x)$ denotes the collection of reduced words in the generators and their inverses that begin with the letter x. Although the family of subsets $\mathscr{W}$ does not exhaust all of F_2 (its union omits the empty word, a.k.a. the identity element of F_2), it can be easily modified to give a more symmetric statement.

Proposition 11.1. *There exists a pairwise disjoint family* $\{E_1, E_2, E_3, E_4\}$ *of subsets of* F_2 *such that*

$$F_2 = E_1 \uplus \ldots \uplus E_4 = E_1 \uplus aE_2 = E_3 \uplus bE_4. \tag{11.1}$$

Proof. The provisions of (11.1) are fulfilled with $E_1 = W(a) \backslash \{a, a^2, a^3, \ldots\}$, $E_2 = W(a^{-1}) \uplus \{e, a, a^2, a^3, \ldots\}$, $E_3 = W(b)$, and $E_4 = W(b^{-1})$. □

This "paradoxical" nature of F_2 has far-reaching consequences. To see how it works, assume that X is an arbitrary set and G a group of self-maps of X. Thus G is a family of self-maps of X that is closed under composition, contains the identity map on X, and contains the (compositional) inverse of each of its members. In particular, each $g \in G$ is a *bijection* of X: a one-to-one mapping taking X onto itself. To say that a set is *partitioned* by a family of subsets means that the subsets of the family are nonempty, pairwise disjoint, and that their union is the whole set. We'll call such a subset family a *partition* of the ambient set.

Definition 11.2 (Paradoxical set)**.** To say that a subset E of X is *G-paradoxical* means that there exist:

(a) A partition of E into finitely many subsets $\{E_1, E_2, \ldots, E_n\}$,
(b) A collection $\{g_1, g_2, \ldots g_n\}$ of elements of G, and
(c) An integer $1 \leq m < n$, such that each family $\{g_j E_j\}_1^m$ and $\{g_j E_j\}_{m+1}^n$ is a partition of E.

Thus "E is G-paradoxical" means that E has a partition whose members can be disjointly reassembled, via transformations in G, into *two* copies of E.

If the group G is understood, we'll abbreviate "G-paradoxical" to just "paradoxical." To say that G itself is paradoxical means that it's paradoxical with respect to the group of left-translation mappings $x \to gx$ $(x, g \in G)$ it induces upon itself. Thus Proposition 11.1 shows that F_2 is paradoxical, with $n = 4$ and $m = 2$ in Definition 11.2. We closed the previous chapter by showing that F_2 is not amenable. In fact this is true of *every* paradoxical group, as the following exercise shows.

Exercise 11.1. Show that: if a set X is paradoxical with respect to a group G of its self-mappings, then $\mathscr{P}(X)$ supports no G-invariant finitely additive probability measure. *Corollary.* No paradoxical group is amenable.

Corollary 11.3. *Neither the closed unit disc* Δ *of* $\mathbb{R}^2$ *nor the unit circle* $\mathbb{T}$ *is paradoxical with respect to the group of rotations about the origin.*

Proof. This follows immediately from Exercise 11.1 above, thanks to Corollary 10.9, which establishes the existence of rotation-invariant finitely additive probability measures on $\mathscr{P}(\Delta)$ and $\mathscr{P}(\mathbb{T})$. □

Exercise 11.2. In Definition 11.2 let's call the set E *paradoxical using n pieces*, or more succinctly: *n-paradoxical*. For example, the free group F_2 is 4-paradoxical.

(a) Show that the ambient set X itself cannot be n-paradoxical for $n < 4$.

(b) Show that a *subset* of X can be n-paradoxical for $n < 4$. (Suggestion: Show that with respect to the bijection group of $\mathbb{Z}$, the subset of natural numbers is 2-paradoxical.)

The next result asserts that paradoxicality can often be transferred from a group to a set upon which that group acts; it is the key to all that follows.

Definition 11.4. To say such a group G of self-maps of a set X is *fixed-point free* on a subset E of X means that no element of G, other than the identity map, can fix a point of E.

Theorem 11.5 (The Transference Theorem)**.** *Suppose X is a set and G a fixed-point free group of self-maps of X. If G is paradoxical, then X is G-paradoxical.*

Proof. We're given: a "replicator family" $\{E_j\}_1^n$ that partitions G, a corresponding family $\{g_j\}_1^n$ of elements of G and an integer m with $1 \le m < n$ such that each "replicant family" $\{g_jE_j\}_{j=1}^m$ and $\{g_jE_j\}_{j=m+1}^n$ also partitions G. We want to show that this situation can be "lifted" to X.

For $x \in X$ the subset $Gx = \{gx : g \in G\}$ is called the *G-orbit of x*. It's easy to check (exercise) that: *The G-orbits partition X*. Consequence: we have $X = \uplus_{m\in M} Gm$ where $M \subset X$ is a "choice set" consisting of one element chosen from each G-orbit.[1] For $g \in G$ let's call the set $gM = \{gm : m \in M\}$ the "co-orbit" of g. The key to transference is the following:

Claim: The co-orbits partition X, i.e., $X = \biguplus\{gM\colon\, g \in G\}$.

Proof of Claim. Observe first that *the co-orbits exhaust X* (proof: $\cup_{g\in G} gM = GM = X$). Thus we need only show that the co-orbits are pairwise disjoint. To this end suppose g and h belong to G and $gM \cap hM \neq \emptyset$. Then there exist points $m_1, m_2 \in M$ such that $gm_1 = hm_2$, so $h^{-1}gm_1 = m_2$, hence m_2 belongs to the G-orbit of m_1. By the definition of our choice set M we must therefore have $m_1 = m_2$, which provides a fixed point for the map $h^{-1}g \in G$. Since G is fixed-point free on X this forces $h^{-1}g$ to be the identity map on X, so $h = g$ and therefore $gM = hM$. Thus, given g and h in M with $g \neq h$, the co-orbits gM and hM must be disjoint, as desired.

For $A \subset G$ let $A^* = AM = \{a(m)\colon\, m \in M, a \in A\}$. Then thanks to the Claim:

If the family of sets $\{A_j\}_1^n$ *partitions G then* $\{A_j^*\}_1^n$ *partitions X.*

Figure 11.1 illustrates the situation.

[1] *Warning:* In general we need the Axiom of Choice (Appendix E.3, p. 209) to do this.

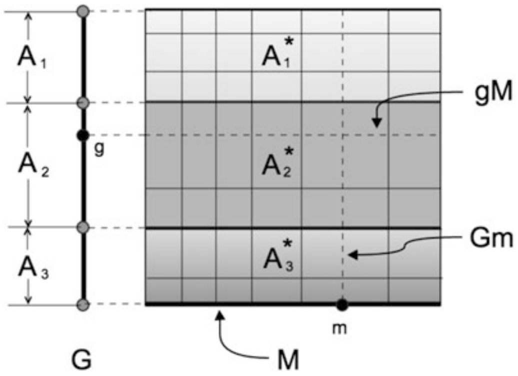

Fig. 11.1 $G = A_1 \uplus A_2 \uplus A_3 \implies X = A_1^* \uplus A_2^* \uplus A_3^*$

Thus our replicator partition $\{E_j\}_1^n$ of G can be transferred to a partition $\{E_j^*\}_1^n$ of X. Similarly the replicant partitions $\{g_jE_j\}_1^m$ and $\{g_jE_j\}_{m+1}^n$ of G transfer to replicant partitions $\{g_jE_j^*\}_1^m$ and $\{g_jE_j^*\}_{m+1}^n$ of X, establishing the G-paradoxicality of X. □

Corollary 11.6. *Each group with a paradoxical subgroup is itself paradoxical.*

Proof. Every subgroup acts freely, by group multiplication, on its parent group. Thus by Theorem 11.5, if the subgroup is paradoxical then so is its parent. □

Exercise 11.3. Show that when G acts freely on X, the family of co-orbits is *transverse* to the family of orbits, i.e., the intersection of each co-orbit with an orbit is a singleton.

Exercise 11.4 (Converse to Theorem 11.5). Suppose G is a group of self-maps of a set X. Show that if X is G-paradoxical, then G is paradoxical. (For this one it's not necessary that G act freely on X.) *Suggestion.* Suppose $\{E_j^*\}_1^n$, $\{g_j\}_1^n$, and $1 \le m < n$ "witness" the G-paradoxicality of X. Fix $x \in X$ and define $E_j = \{g \in G : gx \in E_j^*\}$. Show that the E_j's, $g_j's$, and m witness paradoxicality for G.

Exercise 11.5. For a group G of self-maps of a set X, let C denote the set of points of X, each of which is fixed by some non-identity element of G. Show each map in G takes C, and therefore $X \backslash C$, onto itself. Thus G is a set of self-maps of $X \backslash C$ that is fixed-point free on that set.

11.2 The Hausdorff Paradox

In this section we'll work on the unit sphere S^2 of $\mathbb{R}^3$: the set of points of three dimensional euclidean space that lie at distance 1 from the origin. Let $\mathscr{R}$ denote the group of rotations of $\mathbb{R}^3$ about the origin. For the rest of this chapter we'll treat the notion of "three dimensional rotation" intuitively, taking for granted that each rotation has a "center" through which passes an "axis," every point of which it fixes, and that $\mathscr{R}$ is a group under composition—a group that acts on S^2. All these facts are established in Appendix D (Theorem D.7), where it's shown that $\mathscr{R}$ is isomorphic to the group SO(3) of 3×3 orthogonal matrices with determinant 1.

Theorem 11.7 (The Hausdorff Paradox, c. 1914)**.** *$S^2 \backslash C$ is $\mathscr{R}$-paradoxical for some countable subset C of S^2.*

This result follows quickly from Theorem 11.5 and the following property of the rotation group $\mathscr{R}$, the proof of which we'll defer for just a moment.

Proposition 11.8. *$\mathscr{R}$ contains a free subgroup on two generators.*

What's being asserted here is:

> There exist two rotations $\rho, \sigma \in \mathscr{R}$ with the property that no nonempty reduced word in the "alphabet" $\mathscr{A} = \{\rho, \sigma, \rho^{-1}, \sigma^{-1}\}$ represents the identity transformation.

A "word" in the alphabet $\mathscr{A}$ is a string of symbols $x_1 x_2 \ldots x_n$ with each "letter" x_j an element of $\mathscr{A}$. Each such word "represents" the element of $\mathscr{R}$ obtained by viewing juxtaposition of letters as group multiplication (in this case, composition of mappings). As in the case of F_2, to say a word is "reduced" means that no letter stands next to its inverse.

Granting the above reformulation of the statement of Proposition 11.8, it's fortunate that only one reduced word can represent a given element of $\mathscr{R}$. Equivalently:

> Starting with a word composed of "letters" in the alphabet $\mathscr{A}$, the same reduced word results, no matter how the reduction is performed.

Proof. Suppose $v = x_1 x_2 \ldots x_m$ and $w = y_1 y_2 \ldots y_n$ are two different reduced words in the alphabet $\mathscr{A}$. We wish to prove that they multiply out to different group elements. We may without loss of generality assume that $x_m \neq y_n$ (else cancel these, and keep canceling rightmost letters until you first encounter ones that are distinct; this must happen eventually since $v \neq w$). Let g denote the element of $\mathscr{R}$ you get by interpreting v as a group product, and let $h \in \mathscr{R}$ correspond in this way to w. The word

$$z = vw^{-1} = x_1 x_2 \ldots x_m y_n^{-1} y_{n-1}^{-1} \ldots y_2^{-1} y_1^{-1}$$

corresponds to the group element gh^{-1}.

Claim. *z is a reduced word.*

For this, note that since v and w are reduced, the only cancellation possible in z is at the place where v and w^{-1} join up (w^{-1} is also reduced), i.e., at the pair $x_m y_n^{-1}$. But $x_m \neq y_n$, so no cancellation occurs there, either.

Since z is not the empty word, the property asserted above for the generators ρ and σ guarantees that gh^{-1} is not the identity element of $\mathscr{R}$, i.e., $g \neq h$, so different reduced words in the alphabet $\mathscr{A}$ must correspond to different group elements—as desired. □

Thus we can view the subgroup $\mathscr{F}$ of $\mathscr{R}$ generated by ρ and σ as giving an alternate construction of the free group F_2 on two generators; in particular, it's paradoxical!

Proof of Theorem 11.7. The subgroup $\mathscr{F}$ of $\mathscr{R}$ is countable, and each of its non-identity elements has exactly two fixed points (the points of intersection of the axis of that rotation with S^2). Thus the set C of these fixed points is countable, and by Exercise 11.5, $\mathscr{F}$ is a group of self-maps of $S^2\backslash C$ that is fixed-point free on that set. Thus the Transference Theorem (Theorem 11.5) guarantees that $S^2\backslash C$ is paradoxical for $\mathscr{F}$, and therefore also for $\mathscr{R}$. □

Proof of Proposition 11.8. Choose $\rho \in \mathscr{R}$ to be rotation through $\theta = \sin^{-1}\left(\frac{4}{5}\right)$ radians about the z-axis and σ to be rotation through the same angle about the x-axis. We'll identify these maps with the matrices that represent them relative to the standard unit-vector basis of $\mathbb{R}^3$:

$$\rho = \begin{pmatrix} \frac{3}{5} & -\frac{4}{5} & 0 \\ \frac{4}{5} & \frac{3}{5} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad \sigma = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{3}{5} & -\frac{4}{5} \\ 0 & \frac{4}{5} & \frac{3}{5} \end{pmatrix}.$$

Since ρ and σ are orthogonal matrices their inverses are their transposes, so to say a reduced word of length n in these matrices and their inverses does not multiply out to the identity matrix is to say that the corresponding word in 5ρ, 5σ, and *their transposes* does not multiply out to 5^n times the identity matrix. For *this* it's enough to show that no such word multiplies out to a matrix all of whose entries are divisible by 5, i.e., that *over the field $\mathbb{Z}_5$ of integers modulo 5, no such word multiplies out to the zero-matrix!*

Over the field $\mathbb{Z}_5$ our matrices 5ρ, 5σ, and their transposes become

$$r = \begin{pmatrix} 3 & 1 & 0 \\ 4 & 3 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad r' = \begin{pmatrix} 3 & 4 & 0 \\ 1 & 3 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad s = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 3 & 1 \\ 0 & 4 & 3 \end{pmatrix}, \quad s' = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 3 & 4 \\ 0 & 1 & 3 \end{pmatrix}.$$

Let's call a word in the letters r, r', s, s' *admissible*[2] if r never stands next to r', and s never next to s'.

Our job now is to show that no admissible word in these new matrices multiplies out to the zero-matrix. We'll do this by identifying each matrix with the linear transformation it induces by left-multiplication on the (column) vector space $\mathbb{Z}_5^3$, and proving something more precise:

CLAIM. *The kernel of each admissible word in the letters r, r', s, s' is the kernel of its last letter.*

Proof of Claim. Each of the matrices r, r', s, s' has one dimensional range (i.e., column space) and two dimensional kernel. Upon calculating these ranges and kernels explicitly we find that the ranges of the "r-matrices" intersect the kernels of "s-

[2] We eschew the term "reduced" because, while in our original setup we had, e.g., $\rho\rho^{-1} = \rho^{-1}\rho = I$, now we have $rr' = r'r = 0$.

matrices" in $\{0\}$, and the same is true of the way the kernels of r-matrices intersect the ranges of s-matrices.

Now proceed by induction on word-length. The result is trivial for words of length one. Suppose $n \geq 1$ and that the kernel of each admissible word of length n in r, r', s, s' equals the kernel of that word's last letter. We wish to prove that the same is true of every admissible word of length $n+1$. Let w be such a word, so $w = va$ where v is an admissible word of length n and $a \in \{r, r', s, s'\}$. Then $x \in \ker w$ means that $vax = 0$, i.e., that $ax \in \ker v \cap \operatorname{ran} a$. Since w is an admissible word, the last letter of v, call it b, is not a', and by the induction hypothesis $\ker v = \ker b$. Thus $ax \in \ker b \cap \operatorname{ran} a = \{0\}$, so $x \in \ker a$. We've shown that $\ker w \subset \ker a$. The opposite inclusion is trivial, so $\ker w = \ker a$, as we wished to show. □

Corollary 11.9. *The group $\mathscr{R}$ of rotations of $\mathbb{R}^3$ about the origin is paradoxical, hence not amenable.*

Proof. $\mathscr{R}$ inherits the paradoxicality of its subgroup $\mathscr{F}$ (Corollary 11.6, p. 134), hence it's not amenable (Exercise 11.1, p. 132). □

11.3 Equidecomposability

According to Hausdorff's Paradox, if we remove a certain countable subset from S^2 then what remains is paradoxical with respect to $\mathscr{R}$, the group of rotations of $\mathbb{R}^3$ about the origin. In the next section we'll show, using an "absorption" technique similar the one used to prove Proposition 11.1, that S^2 itself is paradoxical with respect to $\mathscr{R}$. To do this efficiently it will help to have some new terminology.

For the rest of this section, G will denote a group of self-maps of a set X.

Definition 11.10 (Equidecomposability)**.** For subsets E and F of X: To say E is G-*equidecomposable* with F means that there exists a partition $\{E_i\}_1^n$ of E, a partition $\{F_i\}_i^n$ of F, and mappings $\{g_i\}_1^n \subset G$ such that $F_i = g_i E_i$ $(1 \leq i \leq n)$.

Since the inverse of each map in G also belongs to G, it's clear that this notion of "equidecomposable" is symmetric: E is G-equidecomposable with F if and only if F is G-equidecomposable with E. In this case we'll just say "E and F are G-equidecomposable," and use the notation "$E \sim_G F$" to abbreviate the situation. Usually the group G is understood, in which case we'll just say "E and F are equidecomposable," and write $E \sim F$. If we wish to be more precise we'll say "E and F are equidecomposable using n pieces," and write $E \sim_n F$.

The notion of "equidecomposability" allows an efficient restatement of the definition of paradoxicality (Definition 11.2):

Proposition 11.11 (Definition of "Paradoxical" revisited)**.** *A subset E of X is G-paradoxical if and only if there exists a partition of E into subsets A and B such that $A \sim E \sim B$.*

It's an easy exercise to show the relation "$\sim$" on $\mathscr{P}(X)$ is reflexive ($E \sim E$ for every $E \subset X$) and symmetric ($E \sim F \implies F \sim E$). In fact:

Theorem 11.12. *Equidecomposability is an equivalence relation.*

Proof. We need only prove transitivity. Suppose $E \sim F$ and $F \sim H$ for subsets E, F, H of X. Thus there exist partitions $\{E_i\}_1^n$ and $\{F_i\}_1^n$ of E and F, respectively, and transformations $\{g_i\}_1^n \subset G$ such that $g_i E_i = F_i$ for $1 \le i \le n$. There also exist partitions $\{F'_j\}_1^m$ and $\{H_j\}_1^m$ of F and H, respectively, and transformations $\{h_j\}_1^m \subset G$ such that $h_j F'_j = H_j$. Let $E_{i,j} = E_i \cap g_i^{-1}(F_i \cap F'_j)$, and set $\gamma_{i,j} = h_j g_i$ on $E_{i,j}$. Thus each $\gamma_{i,j} \in G$, and one checks easily that (after removing empty sets, if necessary) $\{E_{i,j} : 1 \le i \le n, 1 \le j \le m\}$ and $\{\gamma_{i,j} E_{i,j} : 1 \le i \le n, 1 \le j \le m\}$ partition E and H, respectively. Thus $E \sim H$, as desired. □

The notion of "same cardinality" is defined in terms of arbitrary bijections. In this vein, "equidecomposable" is a refinement of that concept, defined in terms of special bijections. More precisely:

Definition 11.13 (Puzzle Map)**.** For subsets E and F of X, to say a bijection φ of E onto F is a *puzzle map* (more precisely: a "G-puzzle map") means that there is a partition $\{E_i\}_1^n$ of E and transformations $\{g_i\}_1^n \subset G$ such that $\varphi \equiv g_i$ on E_i.

The terminology suggests that we think of E as a jigsaw puzzle assembled from some finite collection of pieces, which the puzzle map φ reassembles into another jigsaw puzzle F. With this definition we have the following equivalent formulation of the notion of equidecomposability:

Proposition 11.14 (Equidecomposability via Puzzle Maps)**.** *Subsets E and F of X are G-equidecomposable if and only if there is a G-puzzle map taking E onto F.*

The fact that G-equidecomposability is an equivalence relation can be explained in terms of puzzle maps: reflexivity means that the identity map is a puzzle map, symmetry means that the inverse of a puzzle map is a puzzle map, and the just-proved transitivity means that compositions of puzzle maps are puzzle maps.

The usefulness of equidecomposability stems from the next result, which asserts that paradoxicality is a property, not just of subsets of X, but actually of $\sim_G$ equivalence classes of subsets.

Corollary 11.15. *Suppose E and F are G-equidecomposable subsets of X. Then E is G-paradoxical if and only if F is G-paradoxical.*

Proof. By symmetry we need only prove one direction. Suppose E is G-paradoxical. Proposition 11.11 provides us with disjoint subsets A and B of E such that $A \sim E \sim B$. Since $E \sim F$ we're given a puzzle map φ mapping E onto F. Since φ is one-to-one, $A' = \varphi(A)$ and $B' = \varphi(B)$ are disjoint subsets of F, and since the restriction of a puzzle map is clearly a puzzle map we know that $A' \sim A$ and $B' \sim B$. Thus by transitivity: $A' \sim A \sim E \sim F$ and $B' \sim B \sim E \sim F$, hence $A' \sim F \sim B'$, so F is G-paradoxical by Proposition 11.11. □

11.4 The Banach–Tarski Paradox for S^2 and $\mathbb{B}^3$

We know so far that if we remove a certain countable subset C from S^2, then the remainder $S^2\backslash C$ is paradoxical for the group $\mathscr{R}$ of rotations of $\mathbb{R}^3$ about the origin. Aided by our work on equidecomposability, we can now give an efficient proof that S^2 itself is paradoxical. For this we'll build on the "absorption" idea that established the paradoxicality of the free group F_2 (Proposition 11.1). Here is the main tool:

Lemma 11.16 (The Absorption Lemma)**.** *Suppose X is a set, E is a subset of X, and C is a countable subset of E. Suppose G is an uncountable group of self-maps of X that takes C into E and is fixed-point free on C. Then E and $E\backslash C$ are G-equidecomposable.*

Proof. The key here is to establish the following:

CLAIM. There exists $g \in G$ such that the family of sets $\{g^n(C) : n \in \mathbb{N} \cup \{0\}\}$ is pairwise disjoint.

Granting this: Let $C_\infty = \biguplus_{n=0}^{\infty} g^n(C)$. Then $C_\infty \subset E$ and, since the sets $g^n(C)$ are pairwise disjoint, $g(C_\infty) = C_\infty\backslash C$. Thus

$$E\backslash C = (E\backslash C_\infty) \uplus (C_\infty\backslash C) = (E\backslash C_\infty) \uplus g(C_\infty) \sim_G (E\backslash C_\infty) \uplus C_\infty = E$$

which establishes the theorem, showing in addition that only two pieces suffice.

Proof of Claim. It's enough to show that for some $g \in G$ we have $g^n(C) \cap C = \emptyset$ for each $n \in \mathbb{N}$. Indeed, once this has been established then given positive integers m and n with $n > m$ we'll have

$$g^n(C) \cap g^m(C) = g^m(g^{n-m}(C) \cap C) = g^m(\emptyset) = \emptyset.$$

Thus to finish the proof it's enough to show that the subset H of G, consisting of maps g for which $g^n(C) \cap C \neq \emptyset$ for some $n \in \mathbb{N}$, is at most countable; the existence of the desired $g \in G$ will then follow from the uncountability of G.

To this end, note that given c and c' in C there is at most one $h \in G$ with $h(c) = c'$. For if $h' \in G$ also takes c to c' then $h^{-1}h'$ fixes c, hence (because the action of G is fixed-point free on C) $h^{-1}h'$ is the identity map on X, i.e., $h = h'$. Now $h \in H$ if and only if there exist points c and c' in C and $n \in \mathbb{N}$ such that $h^n(c) = c'$. By the uniqueness just established, if $k \in G$ has the property that $k^m(c) = c'$ for some non-negative integer m, then $k = h^{n-m}$. Thus given the pair (c, c'), there's at worst a countable family of maps $h \in G$ for which some (integer) power of h takes c to c'. Since there are only countably many such pairs (c, c'), the set of all such maps h, i.e., the set H, is countable. □

Theorem 11.17 (Banach–Tarski for S^2)**.** *The unit sphere S^2 of $\mathbb{R}^3$ is $\mathscr{R}$-paradoxical.*

Proof. We know from the Hausdorff Paradox (Theorem 11.7) that S^2 contains a countable subset C such that $S^2\backslash C$ is paradoxical. Choose a line L through the origin that does not intersect C, and let G denote the subgroup of $\mathscr{R}$ consisting of rotations with axis L. Thus G is an uncountable group that is fixed-point free on C, so the Absorption Lemma with $X = E = S^2$, tells us that S^2 is G-equidecomposable (hence also is $\mathscr{R}$-decomposable) with $S^2 \backslash C$. Corollary 11.15 now guarantees that S^2 inherits the $\mathscr{R}$-paradoxicality of $S^2 \backslash C$. □

Theorem 11.17 gives an embryonic Banach–Tarski Paradox for the closed unit ball $\mathbb{B}^3$, i.e., the set of vectors in $\mathbb{R}^3$ of that lie at distance at most 1 from the origin.

Corollary 11.18. *$\mathbb{B}^3 \backslash \{0\}$ is $\mathscr{R}$-paradoxical.*

Proof. The $\mathscr{R}$-paradoxicality of S^2 means that it contains disjoint subsets A and B such that

$$A \sim S^2 \sim B \tag{11.2}$$

(Proposition 11.11). Let $A^* = \bigcup_{a \in A} \{ra\colon 0 < r \leq 1\}$, and similarly define B^*. Thus $\{A^*, B^*\}$ is a partition of $\mathbb{B}^3 \backslash \{0\}$, and $A^* \sim \mathbb{B}^3 \backslash \{0\} \sim B^*$ via the rotations responsible for (11.2). Thus $\mathbb{B}^3 \backslash \{0\}$ is $\mathscr{R}$-paradoxical. □

Exercise 11.6. Show that both $\mathbb{R}^3 \backslash \mathbb{B}^3$ and $\mathbb{R}^3 \backslash \{0\}$ are $\mathscr{R}$-paradoxical.

The next exercise gives a nontrivial instance of the failure of the Transference Theorem (Theorem 11.5) if the action of the group G is not fixed-point free.

Exercise 11.7. Show that, with respect to the group $\mathscr{R}$ of rotations of $\mathbb{R}^3$ about the origin, $\mathbb{B}^3$ is *not* paradoxical.

Exercise 11.7 also shows that in order to establish the full Banach–Tarski Paradox for $\mathbb{B}^3$ we'll need to go beyond the group of rotations about the origin. Let $\mathscr{G}$ denote the group of rigid motions of $\mathbb{R}^3$ (i.e., the collection of isometric mappings taking $\mathbb{R}^3$ onto itself). In particular, every rotation, whether centered at the origin or not, belongs to $\mathscr{G}$.

Theorem 11.19 (Banach–Tarski for $\mathbb{B}^3$)**.** *The three dimensional unit ball is a $\mathscr{G}$-paradoxical subset of $\mathbb{R}^3$.*

Proof. Let L be the line through the point $(0,0,\frac{1}{2})$ parallel to the x-axis. Let $\mathscr{G}_L$ denote the subgroup of $\mathscr{G}$ consisting of rotations with axis L. Trivially $\mathscr{G}_L$ is fixed-point free on the singleton $\{0\}$, which it takes into $\mathbb{B}^3$. Upon setting $X = \mathbb{R}^3$, $E = \mathbb{B}^3$, and $C = \{0\}$ in the Absorption Lemma we see that $\mathbb{B}^3$ and $\mathbb{B}^3 \backslash \{0\}$ are $\mathscr{G}_L$-equidecomposable, hence $\mathscr{G}$-equidecomposable (using two pieces). Thus $\mathbb{B}^3$ inherits the $\mathscr{G}$-paradoxicality of $\mathbb{B}^3 \backslash \{0\}$. □

Thus each closed ball can be thought of as a three dimensional jigsaw puzzle that can be reassembled, using only rotations (not all of them about the ball's center), into *two* closed balls of the same radius. This raises further questions: Is every ball $\mathscr{G}$-equidecomposable with every other ball? With a cube? We'll take up these matters in the next section.

Exercise 11.8. Continuing in the spirit of Exercise 11.2: in Definition 11.10 let's say that the sets E and F are *n-equidecomposable,* and write $E \sim_n F$.

(a) Show that $E \sim_m F$ and $F \sim_n H$ imply $E \sim_{mn} H$.
(b) Show that S^2 is 8-paradoxical with respect to the rotation group $\mathscr{R}$.
(c) Show that $\mathbb{B}^3$ is 16-paradoxical with respect to the isometry group $\mathscr{G}$.

11.5 Banach–Tarski beyond $\mathbb{B}^3$

Galileo in 1638 discussed the paradox one encounters in comparing the sizes of infinite sets. Using the notation "$A \sim B$" for "there exists a bijection of set A onto set B" (i.e., "A and B have the same cardinality") Galileo's Paradox can be expressed as follows:

> If $\mathbb{N}$ is the set of natural numbers, S the subset of squares, and T the subset of nonsquares, then, even though $\mathbb{N}$ is the disjoint union of S and T, it's nonetheless true that $S \sim \mathbb{N} \sim T$.

Proposition 11.11 phrases the notion of paradoxicality in similar terms, but now using the more sophisticated equivalence relation of "equidecomposability." Like the notion of "same cardinality," equidecomposability can be defined in terms of bijections, but now the bijections are "piecewise congruences," i.e., *puzzle maps* (Proposition 11.14).

The deepest elementary result about "same cardinality" is the Schröder–Bernstein Theorem: if set A has the same cardinality as a subset of set B, and B has the same cardinality as a subset of A, then A and B have the same cardinality. The same is true for equidecomposability; the two results even have a common proof! In this section we'll give this proof and examine its astonishing consequences for the notion of paradoxicality.

We'll assume as usual that G is a group of self-maps of a set X, and we'll continue to write $A \sim_G B$ for "A and B are G-equidecomposable."

Notation 11.20. By "$A \preceq_G B$" we mean "A is G-equidecomposable with a subset of B," i.e., "There is a puzzle map taking A onto a subset of B."

Thus the relation $\preceq_G$ is *reflexive* since the identity map is a puzzle map, and *transitive* since the composition of puzzle maps is a puzzle map. To proceed further we'll need a simple observation about the ordering $\preceq_G$.

Lemma 11.21. *Suppose $\{A_j\}_1^n$ and $\{B_j\}_1^n$ are families of subsets of X, each of which is pairwise disjoint.*

(a) If $A_j \preceq_G B_j$ for each index j, then $\biguplus_{j=1}^n A_j \preceq_G \biguplus_{j=1}^n A_j$.
(b) If $A_j \sim_G B_j$ for each index j, then $\biguplus_{j=1}^n A_j \sim_G \uplus_{j=1}^n A_j$.

Proof. (a) Our hypothesis is that for each j there is a puzzle map φ_j taking A_j into B_j. Then it's easy to check that the map φ defined by setting $\varphi = \varphi_j$ on A_j is a puzzle map taking the union of the A_j's onto the union of the B_j's.

(b) Same as (a), except now the puzzle map φ_j takes A_j *onto* B_j $(1 \le j \le n)$, and therefore φ takes $\biguplus_1^n A_j$ *onto* $\biguplus_1^n B_j$. □

The key to the rest of this section is the fact that $\preceq_G$, in addition to being reflexive and transitive, is also *antisymmetric*, and so induces a partial order on $\mathscr{P}(X)$. This is the content of:

Theorem 11.22 (The Banach–Schröder–Bernstein Theorem)**.** *If A and B are subsets of X with $A \preceq_G B$ and $B \preceq_G A$, then $A \sim_G B$.*

Proof. The hypotheses assert that there are puzzle maps f and g with f taking A onto a subset B_1 of B and g taking B onto a subset A_1 of A. By the Banach Mapping Theorem (Theorem 1.1, p. 8) there is a subset C of A such that g takes $B \backslash f(C)$ *onto* $A \backslash C$. Since g is a puzzle map, and since the restriction of a puzzle map is again a puzzle map, this equation asserts that $B \backslash f(C) \sim A \backslash C$, where here—and in the arguments to follow—we allow ourselves to omit the subscript $\mathscr{G}$. Since f is a puzzle map we know that $f(C) \sim C$. Thus Lemma 11.21 insures that

$$B = (B \backslash f(C)) \cup f(C) \sim (A \backslash C) \cup C = A$$

as desired. □

Previously we noted that for subsets A and B of X:

$$A \subset B \implies A \preceq_G B.$$

Recall the notation $\mathscr{G}$ for the group of all isometric self-maps of $\mathbb{R}^3$.

Corollary 11.23. *Suppose $\{B_j\}_1^n$ is a pairwise disjoint family of subsets of $\mathbb{R}^3$, each of which is $\mathscr{G}$-equidecomposable with $\mathbb{B}^3$. Then $\biguplus_{j=1}^n B_j \sim_{\mathscr{G}} \mathbb{B}^3$.*

Proof. We proceed by induction on n; if $n = 1$ there is nothing to prove, so suppose $n > 1$ and that the result is true for $n-1$. Let $C_1 = \biguplus_{j=1}^{n-1} B_j$ and $C_2 = C_1 \uplus B_n$; our goal is to show that $C_2 \sim \mathbb{B}^3$. Now both C_1 (induction hypothesis) and B_n are $\sim \mathbb{B}^3$ and by the Banach–Tarski Theorem there exists a partition $\{E_1, E_2\}$ of $\mathbb{B}^3$ such that E_1 and E_2 are each $\sim \mathbb{B}^3$. Thus $E_1 \sim C_1$ and $E_2 \sim B_n$, so by Lemma 11.21, $\mathbb{B}^3 = E_1 \uplus E_2 \sim C_1 \uplus B_n = C_2$. □

Corollary 11.24. *Every closed ball in $\mathbb{R}^3$ is $\mathscr{G}$-equidecomposable with every other closed ball.*

Proof. Fix a closed ball B in $\mathbb{R}^3$. It's enough to prove that B is equidecomposable with the closed unit ball $\mathbb{B}^3$.

Suppose first that the radius of B is > 1. Cover B by balls $\{B_j\}_1^n$ of radius equal to one, and "disjointify" this collection of B_j's by setting

$$B'_j = B_j \backslash \cup_{k=j+1}^n B_k.$$

Then $B'_j \subset B_j$ for each index j and the new collection $\{B'_j\}_1^n$ has the same union as the original one; in particular it still covers B. Now let $\{C_j\}_1^n$ be a pairwise disjoint collection of closed balls of radius 1 in $\mathbb{R}^3$. Then

$$\mathbb{B}^3 \preceq B \preceq \biguplus_{j=1}^{n} B'_j \preceq \biguplus_{j=1}^{n} C_j \sim \mathbb{B}^3$$

where the first "inequality" comes from the containment of $\mathbb{B}^3$ in a translate of B, the second one from the containment of B in the union of the B'_j s and the third one from Lemma 11.21 above along with the containment of each B'_j in a translate of the corresponding C_j. Corollary 11.23 provides the final "equality." Thus $\mathbb{B}^3 \preceq B \preceq \mathbb{B}^3$, so $B \sim \mathbb{B}^3$ by the Banach–Schröder–Bernstein Theorem.

If the radius of B is < 1, repeat the above argument with the roles of B and $\mathbb{B}^3$ reversed. If the radius of B is equal to 1 then B, being a translate of $\mathbb{B}^3$, is trivially $\mathscr{G}$-equidecomposable with that set. □

Corollary 11.25 (The "Ultimate" Banach–Tarski Theorem)**.** *Every two bounded subsets of* $\mathbb{R}^3$ *with nonempty interior are* $\mathscr{G}$*-equidecomposable.*

Proof. Let E be a bounded subset of $\mathbb{R}^3$ with nonempty interior. It's enough to show that $E \sim \mathbb{B}^3$. Since E contains a closed ball B we know from Corollary 11.24 that $\mathbb{B}^3 \sim B \preceq E$. Since E is bounded it is contained in a closed ball B', so again by Corollary 11.24: $E \preceq B' \sim \mathbb{B}^3$, hence $\mathbb{B}^3 \preceq E \preceq \mathbb{B}^3$. Thus $E \sim \mathbb{B}^3$ by Banach–Schröder–Bernstein. □

Exercise 11.9 (Paradoxicality revisited). In many expositions of the Banach–Tarski paradox the definition of "paradoxical" is taken to be somewhat less restrictive than the one we've used here (Definition 11.2). Specifically: the "replicator family" $\{E_n\}_1^n$ of that definition is often required only to be pairwise disjoint (not necessarily with union equal to E), while the "replicant families," although still required to exhaust all of E, no longer need to be pairwise disjoint. Show that in this revised definition:

(a) The replicant families can, without loss of generality, be assumed to be pairwise disjoint. Thus we can rephrase the new definition as follows: *There exist disjoint subsets* A_0 *and* B *contained in* E *with* $A_0 \sim E \sim B$.

(b) Let $A = A_0 \uplus E \backslash (A_0 \cap B)$, so that $E = A \uplus B$. Show that $A \sim E$. Thus the new "weakened" definition of paradoxicality is equivalent to the original one.

Notes

A free group of rotations. The idea to consider the two matrices used in the proof of Proposition 11.8, and to transfer the argument to the field $\mathbb{Z}_5$, comes from Terry Tao's intriguing preprint [115].

The Hausdorff Paradox. The original version of Hausdorff's Paradox occurs in [48]; it asserts that there exists a countable subset C of S^2 such that $S^2 \backslash C$ can be partitioned into three subsets A, B, and C such that each is congruent via rotations to the others, *and also to* $B \cup C$. Hausdorff's motivation here was to show that $\mathscr{P}(S^2)$ does not support a rotation-invariant finitely additive probability measure.

References for the Banach–Tarski Paradox. The results of Sects. 11.3–11.5 all come from Banach and Tarski's famous paper [10]. See Chap. 3 of Stan Wagon's book [121] (the gold standard for exposition on the Banach–Tarski Paradox and the research it has inspired up through 1992) for more on the material we've covered here. See also [104, Chap. 1] for another exposition of the Banach–Tarski Paradox and for more recent developments, with the emphasis on amenability. A more popularized exposition of the Banach–Tarski paradox is Leonard Wapner's delightful book [122], which provides much interesting biographical information about the personalities involved, as well as commentary on the foundational issues raised by this amazing theorem.

"n-Paradoxicality." Regarding Exercises 11.2 and 11.8: Raphael Robinson proved in the 1940s that S^2 is 4-paradoxical with respect to the rotation group $\mathscr{R}$ and that $\mathbb{B}^3$ is 5-paradoxical (but not 4-paradoxical) with respect to the full isometry group $\mathscr{G}$. Wagon discusses these matters, with appropriate references, in Chap. 4 of [121].

Amenability and paradoxicality. We've seen that paradoxical groups are not amenable (Exercise 11.1). In the late 1920s Tarski proved that the converse is true: *If a group is not amenable, then it is paradoxical.* See [121, Chap. 9, pp. 125–129] for an exposition of this remarkable theorem.

Galileo's Paradox. In his treatise [41] (pp. 31–33) Galileo observes that the set of squares in $\mathbb{N}$ is in one-to-one correspondence with $\mathbb{N}$ itself, and so has the same size as $\mathbb{N}$. He concludes that size comparisons between infinite sets are impossible.