CAMBRIDGE

Cambridge University Press
978-1-107-04259-9 - The Banach–Tarski Paradox: Second Edition
Grzegorz Tomkowicz and Stan Wagon
Excerpt
More information

PART ONE

Paradoxical Decompositions, or the Nonexistence of Finitely Additive Measures

1

Introduction

It has been known since antiquity that the notion of infinity leads very quickly to seemingly paradoxical constructions, many of which seem to change the size of objects by operations that appear to preserve size. In a famous example, Galileo observed that the set of positive integers can be put into a one-one correspondence with the set of square integers, even though the set of nonsquares, and hence the set of all integers, seems more numerous than the squares. He deduced from this that "the attributes 'equal,' 'greater' and 'less' are not applicable to infinite . . . quantities," anticipating developments in the twentieth century, when paradoxes of this sort were used to prove the nonexistence of certain measures.

An important feature of Galileo's observation is its resemblance to a duplicating machine; his construction shows how, starting with the positive integers, one can produce two sets, each of which has the same size as the set of positive integers. The idea of duplication inherent in this example will be the main object of study in this book. The reason that this concept is so fascinating is that, soon after paradoxes such as Galileo's were being clarified by Cantor's theory of cardinality, it was discovered that even more bizarre duplications could be produced using rigid motions, which *are* distance-preserving (and hence also area-preserving) transformations. We refer to the Banach–Tarski Paradox on duplicating spheres or balls, which is often stated in the following fanciful form: a pea may be taken apart into finitely many pieces that may be rearranged using rotations and translations to form a ball the size of the sun. The fact that the Axiom of Choice is used in the construction makes it quite distant from physical reality, though there are interesting examples that do not need the Axiom of Choice (see Thm. 1.7, §§4.2, 4.3, 11.2).

Two distinct themes arise when considering the refinements and ramifications of the Banach–Tarski Paradox. First is the use of ingenious geometric and algebraic methods to construct paradoxes in situations where they seem impossible and thereby getting proofs of the nonexistence of certain measures. Second, and this comprises Part II of this book, is the construction of measures and their use in showing that some paradoxical decompositions are not possible.

We begin with a formal definition of the idea of duplicating a set using certain transformations. The general theory is much simplified if the transformations used are all bijections of a single set, and the easiest way to do this is to work in the context of group actions. Recall that a group G is said to act on a set X if to each $g \in G$ there corresponds a function (necessarily a bijection) from X to X, also denoted by g, such that for any $g, h \in G$ and $x \in X$, $g(h(x)) = (gh)(x)$ and $e(x) = x$, where e denotes the identity of G.

Definition 1.1. *Let G be a group acting on a set X and suppose $E \subseteq X$ is a nonempty subset of X. Then E is* G-paradoxical *(or paradoxical with respect to G) if, for some positive integers m, n, there are pairwise disjoint subsets $A_1, \ldots, A_n, B_1, \ldots, B_m$ of E and $g_1, \ldots, g_n, h_1, \ldots, h_m \in G$ such that $E = \bigcup g_i(A_i)$ and $E = \bigcup h_j(B_j)$.*

Loosely speaking, the set E has two disjoint subsets ($\bigcup A_i, \bigcup B_j$) each of which can be taken apart and rearranged via G to cover all of E. If E is G-paradoxical, then the sets witnessing that may be chosen so that $\{g_i(A_i)\}$, $\{h_j(B_j)\}$, and $\{A_i\} \cup \{B_j\}$ are each partitions of E. For the first two, one need only replace A_i, B_j by smaller sets to ensure pairwise disjointness of $\{g_i(A_i)\}$ and $\{h_j(B_j)\}$, but the proof that, in addition, $\{A_i\} \cup \{B_j\}$ may be taken to be all of E is more intricate and will be given in Corollary 3.7.

1.1 Examples of Paradoxical Actions

1.1.1 The Banach–Tarski Paradox

Any ball in $\mathbb{R}^3$ is paradoxical with respect to G_3, the group of isometries of $\mathbb{R}^3$.

This result, a paradigm of the whole theory, will be proved in Chapter 3. More generally, we shall consider the possibility of paradoxes when X is a metric space and G is a subgroup of the group of isometries of X (an *isometry* is a bijection from X to X that preserves distance). In the case that G is the group of all isometries of X, we shall suppress G, using simply, E is *paradoxical*. We shall be concerned mostly with the case that X is one of the Euclidean spaces $\mathbb{R}^n$.

1.1.2 Free Non-Abelian Groups

Any group acts naturally on itself by left translation. The question of which groups are paradoxical with respect to this action turns out to be quite fascinating and is discussed in Chapter 12. In this context, the central example is the free group on two generators. Recall that the free group F with generating set M is the group of all finite words using letters from $\{\sigma, \sigma^{-1} : \sigma \in M\}$, where two words are equivalent if one can be transformed to the other by the removal or addition of finite pairs of adjacent letters of the form $\sigma\sigma^{-1}$ or $\sigma^{-1}\sigma$. A word with no such adjacent pairs is called a reduced word, and to avoid the use of equivalence classes, F may be taken to consist of all reduced words, with the group operation being

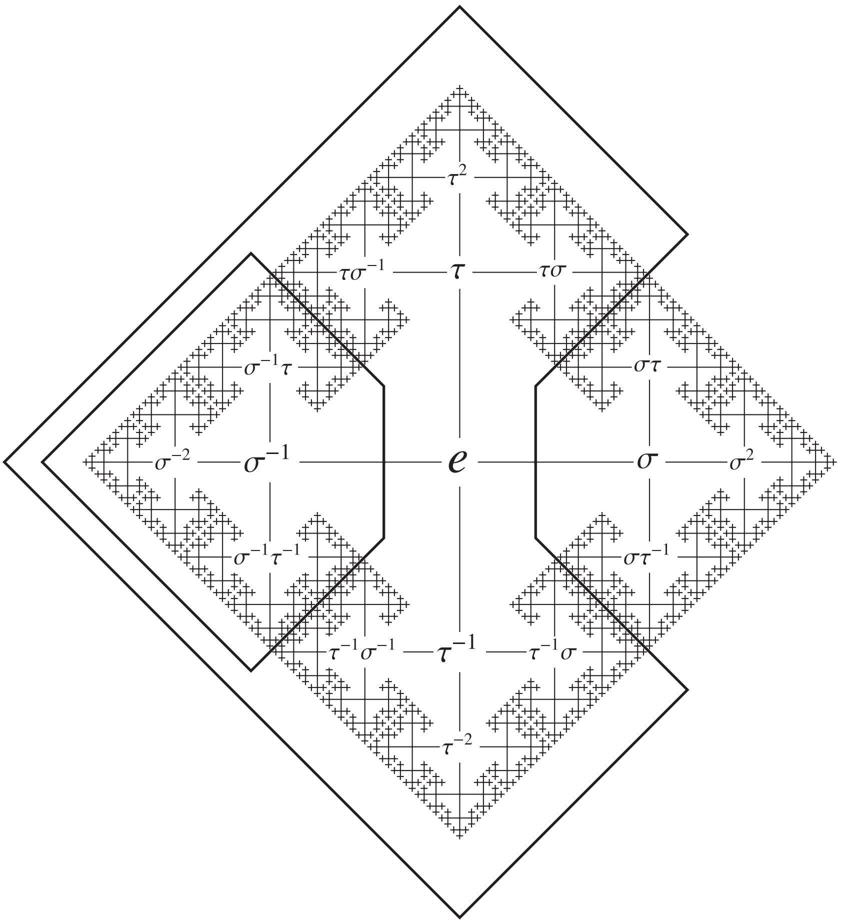

Figure 1.1. The free group of rank 2. The small enclosed region represents $W(\sigma^{-1})$, and left translation of this by σ gives the words in the larger enclosed region. Thus $W(\sigma) \cup \sigma W(\sigma^{-1}) = F$.

concatenation; the concatenation of two words is equivalent to a unique reduced word. (From now on, all words will be assumed to be reduced.) The identity of F, which is denoted by e, is the empty word. A subset S of a group is called *free* if no nonidentity reduced word using elements of S gives the identity. Any two free generating sets for a free group have the same size, which is called the rank of the free group. Free groups of the same rank are isomorphic; any group that is isomorphic to a free group will also be called a free group. See [MKS66] for further details about free groups and their properties.

Theorem 1.2. *A free group F of rank 2 is F-paradoxical, where F acts on itself by left multiplication.*

Proof. Suppose σ, τ are free generators of F. If ρ is one of $\sigma^{\pm 1}, \tau^{\pm 1}$, let $W(\rho)$ be the set of elements of F whose representation as a word in $\sigma, \sigma^{-1}, \tau, \tau^{-1}$ begins, on the left, with ρ. Then $F = \{e\} \cup W(\sigma) \cup W(\sigma^{-1}) \cup W(\tau) \cup W(\tau^{-1})$, and these subsets are pairwise disjoint. Furthermore, $W(\sigma) \cup \sigma W(\sigma^{-1}) = F$ (see Fig. 1.1) and $W(\tau) \cup \tau W(\tau^{-1}) = F$. For if $h \in F \setminus W(\sigma)$, then $\sigma^{-1}h \in W(\sigma^{-1})$ and $h = \sigma(\sigma^{-1}h) \in \sigma W(\sigma^{-1})$. Note that this proof uses only four pieces. □

The preceding proof can be improved so that the four sets in the paradoxical decomposition cover all of F rather than just $F \setminus \{e\}$. The reader might enjoy trying to find such a neat four-piece paradoxical decomposition of a rank 2 free group (or see Fig. 3.2). When we say that a group is paradoxical, we shall be referring to the action of left translation; this should cause no confusion with the usage mentioned in Example 1.1.1.

1.1.3 Free Semigroups

We shall on occasion be interested in the action of a semigroup S (a set with an associative binary operation and an identity) on a set X. Because of the lack of inverses in a semigroup, the function on X induced by some $\sigma \in S$ may not be a bijection; thus it is inappropriate to apply Definition 1.1 to such actions. Nonetheless, there are similarities between free semigroups and free groups, as the following proposition shows. A free semigroup with free generating set T is simply the set of all words using elements of T as letters, with concatenation being the semigroup operation. The rank of a free semigroup is the number of elements in T. A free subsemigroup of a group is a subset of the group that contains e and is closed under the group operation such that the semigroup is isomorphic to a free semigroup.

Proposition 1.3. *A free semigroup S, with free generators σ and τ, contains two disjoint sets A and B such that $\sigma S = A$ and $\tau S = B$. Any group having a free subsemigroup of rank 2 contains a paradoxical set.*

Proof. Let A be the set of words whose leftmost term is σ and B the same using τ. Then $\sigma S = A$ and $\tau S = B$ (see Fig. 1.2). If S is embedded in a group, then S itself is a paradoxical subset of the group because $\sigma^{-1}(A) = S = \tau^{-1}(B)$. □

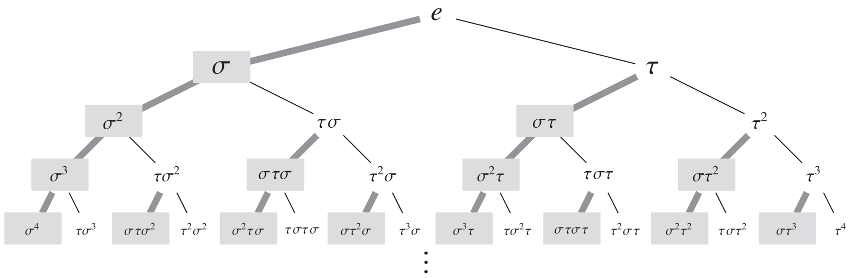

Figure 1.2. A paradox in a group having a free subsemigroup S of rank 2. If A is the set of words with σ on the left (gray background) and B are those with τ on the left (not gray and not e), then $\sigma^{-1}(A) = S = \tau^{-1}(B)$. The thicker edges indicate left multiplication by σ.

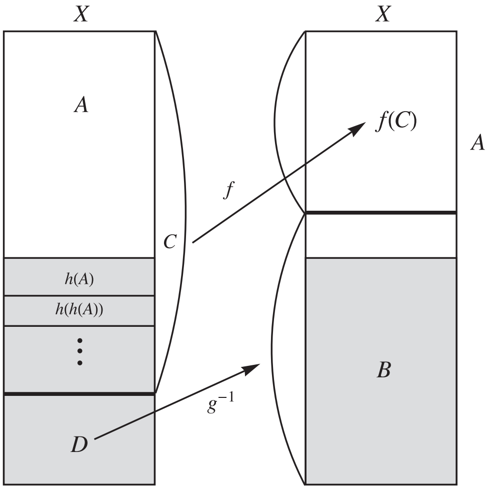

Figure 1.3. The method of constructing the permutation f_1 of X from two bijections f, g from X to subsets of X.

1.1.4 Arbitrary Bijections

The following result, showing that any infinite set is paradoxical using arbitrary bijections, is the modern version of Galileo's observation about the integers. The implications with (c) as hypothesis use the Axiom of Choice (AC). Recall that, in the presence of AC, the cardinality of an infinite set X, $|X|$, is the unique cardinal $\aleph_\alpha$ for which there is a bijection with X. In the absence of AC, $|X|$ is used only in the context of the equivalence relation: $|X| = |Y|$ iff there is a bijection from X to Y.

Theorem 1.4. *The following are equivalent:*

(a) $|X| = 2|X|$.
(b) *X is paradoxical with respect to the group of all permutations of X, that is, all bijections from X to X.*
(c) *X is infinite or empty.*

Proof. We will show (a) $\Rightarrow$ (b) $\Rightarrow$ (c) $\Rightarrow$ (a).

(b) $\Rightarrow$ (c) is clear because finite sets do not admit paradoxes.

(a) $\Rightarrow$ (b). This proof uses the classic back-and-forth idea of the Schröder–Bernstein Theorem (see Thm. 3.6). We start with a partition of X into A and B and bijections $f : X \to A$ and $g : X \to B$. We need bijections f_1 and f_2 from X to X so that f_1 agrees with f on A and f_2 agrees with f on B. To get f_1, let $h = g \circ f : X \to B$ and let $C = A \cup h(A) \cup h(h(A)) \cup \ldots$, which is a disjoint union because h is one-to-one (see Fig. 1.3). Let $D = X \setminus C$. Then f maps C onto $f(C) \subseteq A$, and $g(f(C)) = C \setminus A$. Therefore g is a bijection from $X \setminus f(C)$ to D. So let f_1 be f on C and g^{-1} on D.

The construction of f_2 is similar. Let $h = f \circ f : X \to A$ and $C = B \cup h(B) \cup h(h(B)) \cup \ldots$. Then f maps C to $X \setminus C$ bijectively and f^{-1} is a bijection of $X \setminus C$ with C. So let f_2 be f on C and f^{-1} on $X \setminus C$.

A similar argument gets g_i so that g_1 agrees with g on A and g_2 agrees with g on B.

These functions now give us $X = f_1^{-1}(f(A)) \cup f_2^{-1}(f(B))$. Similarly, X can be realized as a union of permutations restricted to $g(A)$ and $g(B)$. Because $f(A)$, $f(B)$, $g(A)$, $g(B)$ are pairwise disjoint, and this shows that X is paradoxical.

(c) $\Rightarrow$ (a). This is a consequence of the Axiom of Choice. First, it is proved for cardinals by transfinite induction, and then AC (in the form: every set may be mapped bijectively onto a cardinal) is invoked (see [KM68, Chap. 8]). Alternatively, one can give a more direct proof using Zorn's Lemma (see [End77, p. 163]). □

1.2 Geometrical Paradoxes

The first example of a geometrical paradox, that is, one using isometries, arose in connection with the existence of a non–Lebesgue measurable set. The well-known construction of such a set fits into our context if Definition 1.1 is modified to allow countably many pieces. Thus E is *countably G-paradoxical* means that

$$E = \bigcup_{i=1}^{\infty} g_i A_i = \bigcup_{i=1}^{\infty} h_i B_i,$$

where $\{A_1, A_2, \ldots, B_1, B_2, \ldots\}$ is a countable collection of pairwise disjoint subsets of E and $g_i, h_i \in G$. Recall that $\mathbb{S}^1$ denotes the unit circle and $SO_2(\mathbb{R})$ denotes the group of rotations of the circle.

Theorem 1.5 (AC).* *$\mathbb{S}^1$ is countably $SO_2(\mathbb{R})$-paradoxical. If G denotes the group of translations modulo 1 acting on* $[0, 1)$*, then* $[0, 1)$ *is countably G-paradoxical.*

Proof. Let M be a choice set for the equivalence classes of the relation on $\mathbb{S}^1$ given by calling two points equivalent if one is obtainable from the other by a rotation about the origin through a (positive or negative) rational multiple of 2π radians. Because the rationals are countable, these rotations may be enumerated as $\{\rho_i : i = 1, 2, \ldots\}$; let $M_i = \rho_i(M)$. Then $\{M_i\}$ partitions $\mathbb{S}^1$ and, because any two of the M_i are congruent by rotation, the even-indexed of these sets may be (individually) rotated to yield all the M_i, that is, to cover the whole circle. The same is true of $\{M_i : i \text{ odd}\}$. This construction is easily transferred to $[0, 1)$ using the bijection taking $(\cos\theta, \sin\theta)$ to $\theta/2\pi$, which induces an isomorphism of $SO_2(\mathbb{R})$ with G. □

Corollary 1.6 (AC). *(a) There is no countably additive, rotation-invariant measure of total measure 1, defined for all subsets of $\mathbb{S}^1$.*

* In the sequel, theorems whose proof uses the Axiom of Choice will be followed by (AC).

(b) There is a subset of $[0, 1]$ *that is not Lebesgue measurable.*

*(c) There is no countably additive, translation-invariant measure** *defined on all subsets of* $\mathbb{R}^n$ *and normalizing* $[0, 1]^n$.

Proof. (a) Suppose μ is such a measure and let A and B be disjoint subsets of the circle that witness the paradox of Theorem 1.5; then the properties of μ give $1 \geq \mu(A \cup B) = \mu(A) + \mu(B) = 2$, a contradiction.

(b) This follows from (c); in fact, $\{\alpha \in [0, 1) : (\cos\alpha, \sin\alpha) \in M\}$ is not Lebesgue measurable.

(c) For $\mathbb{R}^1$, such a measure cannot exist because its restriction to subsets of $[0, 1]$ would be invariant under translations modulo 1, contradicting Theorem 1.5. Such a measure in $\mathbb{R}^n$ would induce one on the subsets of $\mathbb{R}$, by the correspondence $A \leftrightarrow A \times [0, 1]^{n-1}$. □

The connection between the Axiom of Choice and the existence of nonmeasurable sets is complex, involving the theory of large cardinals and forcing—two branches of contemporary set theory. We consider these connections in more detail in Chapter 15. For now, we note only that (without assuming Choice) the following two assertions are *not* equivalent:

- All sets of reals are Lebesgue measurable.
- There is a countably additive, translation-invariant extension of Lebesgue measure to all sets of reals.

It is known that the second assertion does not imply the first.

It comes as a bit of a surprise that even with the restriction to finitely many pieces, paradoxes can be constructed using isometries. The following construction, the first of its kind, does not require any form of the Axiom of Choice, which adds some weight to the comment of Eves [Eve63] that the result is "contrary to the dictates of common sense." Recall that when no group is explicitly mentioned, it is understood that the isometry group is being used.

Theorem 1.7 (Sierpiński–Mazurkiewicz Paradox). *There is a paradoxical subset of the plane* $\mathbb{R}^2$.

The reason this paradox exists is that the planar isometry group G_2 has a free non-Abelian subsemigroup that acts in a particularly nice way (Thm. 1.8). The single most important idea in constructing a paradoxical decomposition is the transfer of an algebraic paradox from a group or semigroup (as in Prop. 1.3) to a set on which it acts. This technique was first used, independently, by Hausdorff and by Sierpiński and Mazurkiewicz. The next theorem shows that a free subsemigroup exists for plane isometries.

* Measures are allowed to have values in $[0, \infty]$.

Theorem 1.8. *There are two isometries, σ, τ, of $\mathbb{R}^2$ that generate a free subsemigroup of G_2. Moreover, σ and τ can be chosen so that for any two words w_1 and w_2 in σ, τ having leftmost terms σ, τ, respectively, $w_1(0,0) \neq w_2(0,0)$.*

Proof. Choose θ so that $\beta = e^{i\theta}$ is transcendental; $\theta = 1$ works, but it is simpler just to use the fact that the unit circle is uncountable whereas the set of algebraic numbers is countable. Then let σ be rotation by θ and let τ be translation by $(1, 0)$. In $\mathbb{C}$, σ is multiplication by β and τ is addition of 1. We need only prove that σ and τ satisfy the second assertion, because freeness follows from that. For if $w_1 = w_2$, where w_1 and w_2 are distinct semigroup words and one of them is (the identity or) an initial segment of the other, then left cancellation yields $v = e$ for a nontrivial word v. If v has σ on the left, then $v\tau(0) = \tau(0)$, and if v has τ on the left, then $v\sigma(0) = \sigma(0)$, contradicting the second assertion in either case. And if neither is an initial segment of the other, then left cancellation yields w_1 and w_2, which are equal in G_2 but have different leftmost terms.

So, suppose $w_1 = \tau^{j_1}\rho^{j_2}\cdots\tau^{j_m}$ and $w_2 = \rho^{k_1}\tau^{k_2}\cdots\tau^{k_\ell}$, where $m, \ell \geqslant 1$ and each exponent is a positive integer; because $\rho(0) = 0$, it is all right to assume that w_1 and w_2 both end in a power of τ, unless w_2 is simply ρ^{k_1}. Then

$$w_1(0) = j_1 + j_3u^{j_2} + j_5u^{j_2+j_4} + \cdots + j_mu^{j_2+j_4+\cdots+j_{m-1}}$$

and

$$w_2(0) = k_2u^{k_1} + k_4u^{k_1+k_3} + \cdots + k_\ell u^{k_1+k_3+\cdots+k_{\ell-1}} \ (= 0 \text{ if } w_2 = \rho^{k_1}).$$

If $w_1(0) = w_2(0)$, these two expressions may be subtracted to yield a nonconstant polynomial with integer coefficients that vanishes for the value $e^{i\theta}$, and this contradicts the choice of θ. □

Using the isometries of Theorem 1.8 (and working in $\mathbb{C}$), we can prove Theorem 1.7 by directly constructing a paradoxical set in the plane. Let E be the orbit of 0 under the free subsemigroup of Theorem 1.8. Then let $A = \sigma(E)$ and $B = \tau(E)$. Figure 1.4 shows the orbit of 0 in $\mathbb{C}$, where σ is replaced by multiplication by β and τ by addition of 1. The framed numbers form A, and the others are B; we have $E = A/\beta = B - 1$.

Another way of saying this is that E is the set of complex numbers of the form $a_0 + a_1\beta + \cdots + a_n\beta^n$ where n and the coefficients are nonnegative integers. Then A is the set of such numbers for which $a_0 = 0$, and B consists of the others.

We can state this construction in a more abstract form as follows.

Proposition 1.9. *Suppose a group G acting on X contains σ, τ such that for some $x \in X$, any two words in σ, τ beginning with σ, τ, respectively, disagree when applied to x. Then there is a nonempty G-paradoxical subset of X.*

Proof. Let S be the subsemigroup of G generated by τ and ρ, and let E be the S-orbit of x. Then $E \supseteq \tau(E), \rho(E)$, and the hypothesis on x implies that