The Banach–Tarski Paradox

Andrew McKee

November 2020

Abstract

The Banach–Tarski paradox is often stated as follows: given a solid ball in three dimensions it is possible to cut the ball into a finite number of pieces and rearrange these pieces to make two balls, each identical to the original. The result was proved by Banach and Tarski (1924), building on earlier work of Hausdorff (1914). The paradox, along with the fact that no such paradox exists in one or two dimensions, hints at the subtle nature of the concept of volume as well as deep properties of the group of translations and rotations of three-dimensional space.

In these notes we develop the background material and explore some earlier paradoxes, before proving the Banach–Tarski paradox. The final part of the course will discuss how the Banach–Tarski paradox is related to the problem of defining a notion of volume which matches our intuition.

Introduction

How can one give a rigorous definition of volume which matches with our intuition for how volume behaves? More formally: is it possible to define a notion of "volume" for all subsets of $\mathbb{R}^n$ which is invariant under translation and rotation, gives the unit cube a volume of one, and such that the volume of the union of disjoint sets is the sum of the volumes of the individual sets?

In 1901 Lebesgue [4] introduced a way of defining the volume of subsets of $\mathbb{R}^n$, now called *Lebesgue measure*, which satisfies most of the required properties. However, Vitali [8] discovered in 1905 that not every set has a well-defined Lebesgue measure, and his construction showed that there is only hope for a positive answer if one restricts to finite unions of sets. Thus the question remained: can one define a measure on every subset of $\mathbb{R}^n$ which satisfies the properties mentioned above (and extends Lebesgue measure)?

To show that such measures cannot exist mathematicians, such as Hausdorff [3], discovered paradoxes — cutting shapes in to pieces and moving those pieces with rigid motions to form new shapes with a different volume to the original. The most striking of these paradoxes was published in 1924 by Banach and Tarski [1], which is often stated in the form: it is possible to take a solid ball in $\mathbb{R}^3$, divide the ball in finitely many pieces, and move those pieces using only rigid motions to form two solid balls, each identical to the original ball. These results are called paradoxes because only rigid motions are used, and intuition suggests that rigid motions should preserve volume.

The impact of these discoveries is far-reaching. The Banach–Tarski paradox solves the problem above about defining measures in $\mathbb{R}^3$ (and the same idea works for $\mathbb{R}^n$ when $n \geq 3$), and the techniques involved in proving the Banach–Tarski paradox led von Neumann to introduce the notion of *amenability* [9], now an important notion in many areas of mathematics.

Information

These notes formed the basis of a course at the University of Białystok in November–December 2020, based on earlier notes made while supervising a Bachelor's project at Chalmers University of Technology and the University of Gothenburg.

The course was delivered in 30 hours of lectures and 30 hours of problem classes. Some necessary results were given as exercises during lectures, to be attempted by students and discussed during the problem classes; here the exercises are followed by sample solutions.

The following are suggested as references and/or further reading.

- Wagon [10] Updated version of Wagon's comprehensive book on the topic.

- Weston [11] Notes online containing many explicit computations.

- Cohn [2] Measure theory text; see Appendix G for a discussion of the Banach–Tarski paradox.

- Runde [5] Contains an introductory section on the Banach–Tarski paradox.

Contents

Chapter 1

Groups and actions

We assume that the reader is familiar with basic group theory; many references are available for the reader who lacks this background, for example [6].

1.1 Examples

The following examples of groups are important for us later.

Example 1.1.1. *The collection of all real, invertible* $n \times n$ *matrices is a group under matrix multiplication, called the* general linear group (of degree n)*, and denoted* $\mathrm{GL}(n, \mathbb{R})$*; equivalently* $\mathrm{GL}(n, \mathbb{R}) := \{T \in \mathbb{R}^{n\times n} : \det(T) \neq 0\}$.

A basis $\{e_1, \ldots, e_m\}$ of a subspace V of $\mathbb{R}^n$ is said to be *orthonormal* if $\langle e_i, e_j\rangle = 0$ when $i \neq j$ and $\langle e_i, e_i\rangle = 1$, where $\langle\cdot,\cdot\rangle$ denotes the usual dot product on $\mathbb{R}^n$.

Exercise 1.1.2. *Let* V *be a subspace of* $\mathbb{R}^n$ *with orthonormal basis* $B = \{e_1, \ldots, e_m\}$*, and let* $T : V \to V$ *be a linear operator, with matrix* T_B *relative to the orthonormal basis* B*. Show that the following are equivalent:*

i. $\langle Tx, Ty\rangle = \langle x, y\rangle$ *for all* $x, y \in V$*;*

ii. the columns (or rows) of T_B *are mutually orthogonal;*

iii. $T_B^t T_B = I_n$.

Such an operator is called orthogonal.

Solution. (i) $\Longrightarrow$ (ii) In particular we have that $\langle Te_i, Te_j\rangle = \langle e_i, e_j\rangle$ for $1 \leq i, j \leq m$. The vector Te_i is simply the ith column of T_B, so (ii) follows. For the rows apply the same argument to the equivalent condition: $\langle x, y\rangle = \left\langle T^t x, T^t y\right\rangle$ for all $x, y \in V$.

(ii) $\Longleftrightarrow$ (iii) The (i, j) entry of $T_B^t T_B$ is the dot product of column i and column j of the matrix T_B.

(iii) $\Longrightarrow$ (i) For any $x, y \in V$ calculate

$$\langle x, y\rangle = \langle Ix, y\rangle = \left\langle T_B^t T_B x, y\right\rangle = \langle T_B x, T_B y\rangle .$$

□

Example 1.1.3. *The* special orthogonal group (of degree n) *is the subgroup* $\mathrm{SO}(n, \mathbb{R})$ *of* $\mathrm{GL}(n, \mathbb{R})$ *given by*

$$\mathrm{SO}(n, \mathbb{R}) := \{T \in \mathrm{GL}(n, \mathbb{R}) : \det(T) = 1 \textit{ and } T \textit{ is orthogonal}\}.$$

Exercise 1.1.4. *Show that* $\mathrm{SO}(n, \mathbb{R})$ *is a group.*

Solution. It is obvious that $I_n \in \mathrm{SO}(n, \mathbb{R})$. If $T \in \mathrm{SO}(n, \mathbb{R})$ then $T^{-1} = T^t$, and

$$(T^t)^t T^t = TT^t = I_n^t = I_n,$$

so $T^{-1} \in \mathrm{SO}(n, \mathbb{R})$. Finally, if $S, T \in \mathrm{SO}(n, \mathbb{R})$ then

$$(ST)^t(ST) = T^t S^t ST = T^t I_n T = T^t T = I_n,$$

so $ST \in \mathrm{SO}(n, \mathbb{R})$. □

We will see below that the elements of these groups can be viewed as "moving" shapes in $\mathbb{R}^n$. In particular, we will show that $\mathrm{SO}(3, \mathbb{R})$ represents the rotations of the unit ball that will be used in the Banach–Tarski paradox. For now, let us try to develop some intuition for how these matrix groups act.

Exercises 1.1.5. *The introduction states that the Banach–Tarski paradox uses "rigid motions" (i.e. distance-preserving maps) to rearrange the pieces of the unit ball.*

i. What rigid motions are not represented by elements of $\mathrm{SO}(n, \mathbb{R})$*?*

ii. What rigid motions are not represented by elements of $\mathrm{GL}(n, \mathbb{R})$*?*

iii. Does the collection of all rigid motions form a group?

Solution. (i) Reflections correspond to matrices with a negative determinant. These will not be used in our proof of the Banach–Tarski paradox.

(ii) Translations $x \mapsto x + b$ $(x, b \in \mathbb{R}^n)$ do not fix the origin, so do not belong to $\mathrm{GL}(n, \mathbb{R})$ (unless $b = 0$, in which case the translation by b is the identity map I_n).

(iii) Yes. The *Euclidean group (of order n)* is the group containing all reflections, rotations and translations of $\mathbb{R}^n$ and all finite combinations of such. □

We will not use reflections in our proof of the Banach–Tarski paradox, working in a subgroup of the Euclidean group (which contains all rigid motions) called the *Euclidean motion group* which does not contain reflections.

Next we introduce free groups, which will play an important role in the Banach–Tarski paradox. Later we will work with free groups as words on a generating set, so this is how we define them; see [6, Chapter 6] for alternative descriptions.

Example 1.1.6. *Let $S = \{a_1, \ldots, a_n\}$ be a set with n elements, and write $S^{-1} := \{a_1^{-1}, \ldots, a_n^{-1}\}$ for the set of formal inverses of elements of S. A* word *on S is a finite product $s_1 s_2 \cdots s_m$ $(m \geq 1)$, where $s_i \in S \cup S^{-1}$; a* reduced word *on S is a finite product $s_1 s_2 \cdots s_m$ $(m \geq 1)$ such that s_i is never adjacent to its inverse. The* free group on n generators*, denoted $\mathbb{F}_n$, is defined to be the group of all reduced words on $S \cup S^{-1}$, with the group operation given by concatenation and reduction: if w_1 and w_2 are reduced words then their product is the reduced word obtained from $w_1 w_2$; the identity element is the empty word, denoted e, and the inverse of $s \in S$ is $s^{-1} \in S^{-1}$. Similarly one defines the free group $\mathbb{F}_\infty$, by taking S to be a countable set.*

Though we have used $S = \{a_1, \ldots, a_n\}$ in the definition of free groups of arbitrary degree it is conventional to write a and b for the free generators of $\mathbb{F}_2$.

Exercises 1.1.7. *i. Give an example of two pairs of reduced words v_1, v_2 and w_1, w_2 in $\mathbb{F}_3$ such that $v_1 v_2 = w_1 w_2$.*

ii. Show that having the same reduced word is an equivalence relation on the collection of words on $S = \{a_1, \ldots, a_n\}$. Deduce that the group operations on $\mathbb{F}_n$ are well-defined.

iii. Which familiar group is isomorphic to $\mathbb{F}_1$?

iv. Explain why $\mathbb{F}_n$ is countable for any $n \in \mathbb{N}$.

v. *Let m, n be natural numbers with $2 \leq m \leq n$. Define an injective group homomorphism from $\mathbb{F}_m$ to $\mathbb{F}_n$. Can you find an injective group homomorphism going the other way (i.e. can we identify $\mathbb{F}_n$ with a subgroup of $\mathbb{F}_m$)?*

Solution. (i) Take $v_1 = a, v_2 = b, w_1 = abc^{-1}, w_2 = c$, so $v_1 v_2 = ab = w_1 w_2$.

(ii) This exercise is intended to show that it is difficult to define free groups directly using words and reduced words! That the stated property is reflexive, symmetric and transitive follows easily because equality has these properties. Each equivalence class of this relation contains exactly one reduced word. If u, v, w are reduced words such that no cancellations are possible in the concatenations uv and vw then it is obvious that $u(vw) = (uv)w$; if there are cancellations then it is tedious to check! If x, y are reduced words and $x' \in [x]$, $y' \in [y]$ then $[x'y'] = [xy]$ can be seen by first doing the cancellations internal to x' and y', so the product is well-defined. If v is a reduced word $v = s_1 s_2 \cdots s_m$ then $v^{-1} = s_m^{-1} \cdots s_2^{-1} s_1^{-1}$, and it is clear that v^{-1} is also a reduced word with any word $w \in [v^{-1}]$ satisfying $[wv] = [e] = [vw]$.

(iii) Elements of $\mathbb{F}_1$ have the form $aa \cdots a$ or $a^{-1}a^{-1} \cdots a^{-1}$, which we write as a^n and a^{-m} respectively. For $k, l \in \mathbb{Z}$ it is clear that $(a^k)(a^l) = a^{k+l}$, so the map $a^n \mapsto n$ defines an isomorphism from $\mathbb{F}_1$ to $\mathbb{Z}$.

(iv) Define the length of a reduced word to be the total number of generators it contains, including multiplicity. Let A_m denote the set of reduced words of length m, which is finite for $n \in \mathbb{N}$ (and at most countable in the case $\mathbb{F}_\infty$). Clearly $\mathbb{F}_n = \cup_{m=0}^{\infty} A_m$, so it is a countable union of countable sets.

(iv) Let $a_1, \ldots, a_m$ be the generators of $\mathbb{F}_m$ and $b_1, \ldots, b_n$ be the generators of $\mathbb{F}_n$. The map $a_i \mapsto b_i$ clearly extends to an injective homomorphism. For the second part it suffices (from the first part) to find a subgroup of $\mathbb{F}_2$ which is isomorphic to $\mathbb{F}_\infty$. Define $c_n := b^n a b^{-n}$ and convince yourself that the subgroup of $\mathbb{F}_2$ generated by $S := \{c_n : n \in \mathbb{N}\}$ is isomorphic to $\mathbb{F}_\infty$. (In other words, show that there are no unexpected relations between the c_n.) □

The final part of the above exercise shows one of the non-intuitive properties of free groups; we will exploit another strange property of $\mathbb{F}_2$ later to derive the Banach–Tarski Paradox.

1.2 Group actions

Many groups arise naturally as collections of invertible maps on some other object. This concept is formalised as a group action.

Definition 1.2.1. *Let G be a group and X a set. We say G* acts on *X if there is a map $\cdot : G \times X \to X$ such that:*

i. $e \cdot x = x$ *for all* $x \in X$;

ii. $g \cdot (h \cdot x) = (gh) \cdot x$ *for all* $g, h \in G$ *and all* $x \in X$.

Exercise 1.2.2. *Let G be a group. Give an example of an action of G on itself.*

Solution. The left multiplication action: define $\cdot : G \times G \to G$ by $g \cdot h := gh$. That this is an action of G follows directly from the group axioms.

More generally, fix any $r \in G$ and define $g \cdot h := rgr^{-1}h$. Then, for all $h \in G$

$$e \cdot h = rer^{-1}h = rr^{-1}h = h$$

and for $f, g, h \in G$

$$f \cdot (g \cdot h) = rfr^{-1}(g \cdot h) = rfr^{-1}(rgr^{-1}h) = rfgr^{-1}h = (fg) \cdot h,$$

as required. The first example is the special case $r = e$ of the second example.

For a third example define the conjugation action $\cdot : G \times G \to G$ by $g \cdot h := ghg^{-1}$. Then, for all $h \in G$, $e \cdot h = ehe^{-1} = h$, and for $f, g, h \in G$

$$f \cdot (g \cdot h) = f(g \cdot h)f^{-1} = f(ghg^{-1})f^{-1} = (fg)h(fg)^{-1} = (fg) \cdot h.$$

□

Example 1.2.3. *Let n be a natural number and $\mathrm{GL}(n, \mathbb{R})$ the corresponding general linear group. Define $\cdot : \mathrm{GL}(n, \mathbb{R}) \times \mathbb{R}^n \to \mathbb{R}^n$ to be the usual matrix-vector multiplication. This is an action of $\mathrm{GL}(n, \mathbb{R})$ on $\mathbb{R}^n$ because of the properties of matrix multiplication.*

Since $\mathrm{SO}(3, \mathbb{R})$ is a subgroup of $\mathrm{GL}(3, \mathbb{R})$ it follows from Example 1.2.3 above that $\mathrm{SO}(3, \mathbb{R})$ acts on $\mathbb{R}^3$. This action is very important for the Banach–Tarski paradox, so we study it further now.

Exercise 1.2.4. *Let $\mathcal{S}^n$ be the unit sphere in $\mathbb{R}^n$, that is*

$$\mathcal{S}^n := \{x \in \mathbb{R}^n : \|x\| = 1\},$$

where $\| \cdot \|$ denotes the Euclidean distance in $\mathbb{R}^n$.[1] *Show that the restriction of the action of $\mathrm{SO}(n, \mathbb{R})$ on $\mathbb{R}^n$ to $\mathcal{S}^n$ is an action of $\mathrm{SO}(n, \mathbb{R})$ on $\mathcal{S}^n$.*

[1] It is common to denote the unit sphere in $\mathbb{R}^n$ by $\mathcal{S}^{n-1}$, but we prefer to use the convention with $\mathrm{SO}(n, \mathbb{R})$ acting on $\mathcal{S}^n$.

Solution. We only need to check that if $T \in \mathrm{SO}(n, \mathbb{R})$ and $x \in \mathcal{S}^n$ then $Tx \in \mathcal{S}^n$, as the rest of the properties will follow from Example 1.2.3. We know that

$$\|Tx\|^2 = \langle Tx, Tx \rangle = \langle x, x \rangle = \|x\|^2,$$

so if $x \in \mathcal{S}^n$ then so does Tx. □

Exercise 1.2.5. *Let X be a set and G a group acting on X. Define a relation $\sim$ on X by*

$$x \sim y \iff y = g \cdot x \text{ for some } g \in G.$$

Show that this relation is an equivalence relation. The equivalence classes of this relation are called the orbits *of the action.*

Solution. Since $x = e \cdot x$ for all $x \in X$ the relation is reflexive. For symmetry observe that if $x, y \in X$ are such that $y = g \cdot x$ for some $g \in G$ then $x = g^{-1} \cdot y$, so the relation is symmetric. For transitivity, suppose $x, y, z \in X$ with $y = g \cdot x$ and $z = h \cdot y$ for some $g, h \in G$. Then

$$z = h \cdot y = h \cdot (g \cdot x) = (hg) \cdot x,$$

so the relation is transitive. □

Our final aim in this section is to prove that the group which acts on $\mathbb{R}^3$ by rotations about some line through the origin is the group $\mathrm{SO}(3, \mathbb{R})$. The development of these results is based on [2, Appendix G].

Exercise 1.2.6. *Let T be an orthogonal operator on $\mathbb{R}^n$.*

(i) Show that $\det(T)$ *is* 1 *or* -1.

(ii) Show that every real eigenvalue of T has absolute value 1.

(iii) Suppose that $n = 3$. Show that T has at least one real eigenvalue.

Solution. (i) Since $1 = \det(I_n) = \det(T^t T) = \det(T^t)\det(T) = (\det(T))^2$.

(ii) Let $x \neq 0$ be an eigenvector of T with real eigenvalue λ. Then

$$\langle x, x \rangle = \langle Tx, Tx \rangle = \langle \lambda x, \lambda x \rangle = |\lambda|^2 \langle x, x \rangle,$$

so $|\lambda|^2 = 1$.

(iii) If λ is an eigenvalue of T then so is its complex conjugate $\overline{\lambda}$. Since T has three eigenvalues one of them must satisfy $\lambda = \overline{\lambda}$, so is real. □

Lemma 1.2.7. *Let T be an orthogonal operator on $\mathbb{R}^n$ with (real) eigenvalue λ and corresponding eigenvector x. Define $x^\perp := \{y \in \mathbb{R}^n : \langle x, y\rangle = 0\}$.*

(i) The set $x^\perp$ is a subspace of $\mathbb{R}^n$ and $Tx^\perp \subseteq x^\perp$.

(ii) The restriction $T_{x^\perp}$ of T to $x^\perp$ is an orthogonal operator which satisfies $\det(T) = \lambda \det(T_{x^\perp})$.

Proof. (i) Let $y, z \in x^\perp$ and $\mu \in \mathbb{R}$, so

$$\langle x, \mu y + x\rangle = \mu \langle x, y\rangle + \langle x, z\rangle = 0.$$

For the second part take $y \in x^\perp$ and calculate

$$\lambda \langle x, Ty\rangle = \langle \lambda x, Ty\rangle = \langle Tx, Ty\rangle = \langle x, y\rangle = 0,$$

so as $\lambda \neq 0$ we have $\langle x, Ty\rangle = 0$, which means $Ty \in x^\perp$.

(ii) Let e_1 be a normalised vector parallel to x, and extend to a basis B of $\mathbb{R}^n$ containing e_1. Then the first row of T_B is $(\lambda, 0, 0, \ldots, 0)$ and the first column has the same pattern. It follows that the rows of T_B corresponding to $T_{x^\perp}$ (the block obtained by removing the first row and column) are still orthogonal.

The statement about determinants is a general fact. $\square$

In the following exercise you deduce that if $T \in \mathrm{SO}(3, \mathbb{R})$ has an eigenvalue -1 then T is a rotation.

Exercise 1.2.8. *(i) Let S be an orthogonal operator on $\mathbb{R}^2$ with $\det(S) = -1$. Show that both 1 and -1 are eigenvalues of S.*

(ii) Let T be an orthogonal operator on $\mathbb{R}^3$ with $\det(T) = 1$ and an eigenvalue -1. Show that the eigenvalues of T are -1 (multiplicity two) and 1 (multiplicity one).

(iii) Deduce that if T is an orthogonal operator with $\det(T) = 1$ and an eigenvalue -1 then T is rotation by π about some line through the origin.

Solution. (i) Fix an orthonormal basis B of $\mathbb{R}^2$ and write $S_B = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Since $\det(S) = -1$ we have $ad - bc = -1$, while $SS^t = I_n$ gives

$$a^2 + b^2 = 1, \quad c^2 + d^2 = 1, \quad ac + bd = 0.$$

Multiplying the third of these above by c, and substituting for c^2 using the second, gives

$$ac^2 + bcd = 0 \implies a(1-d^2) = -bcd = -d(1+ad),$$

where we used the expression from $\det(S)$ in the final equality. Hence

$$a - ad^2 = -d - ad^2 \implies a + d = 0.$$

Now we calculate the eigenvalues of S, using the above equality:

$$\begin{aligned}\det\begin{pmatrix} a-\lambda & b \\ c & d-\lambda \end{pmatrix} &= (a-\lambda)(d-\lambda) - bc = \lambda^2 - (a+d)\lambda + ad - bc \\ &= \lambda^2 - 1.\end{aligned}$$

Thus $\lambda = \pm 1$.

(ii) Let e_1 be the (normalised) eigenvector corresponding to the eigenvalue -1 and consider $T_{e_1^\perp}$. By Lemma 1.2.7 $1 = -\det(T_{e_1^\perp})$, so $T_{e_1^\perp}$ satisfies the hypotheses of part (i) of this exercise. The statement follows.

(iii) Let $B = \{e_1, e_2, e_3\}$ be an orthonormal basis corresponding to the eigenvalues of T: $Te_1 = e_1$, $Te_2 = -e_2$ and $Te_3 = -e_3$. The matrix of T with respect to this basis is

$$T_B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix},$$

which we recognise as a rotation by π about the axis defined by e_1 and the origin. $\square$

It remains to investigate the case when T does not have -1 as an eigenvalue.

Exercise 1.2.9. *(i) Let S be an orthogonal operator on $\mathbb{R}^2$ with $\det(S) = 1$. Show that for any orthonormal basis B of $\mathbb{R}^2$ there are $a, b \in \mathbb{R}$ with $a^2 + b^2 = 1$, so that $S_B = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$.*

(ii) Deduce that if T is an orthogonal operator on $\mathbb{R}^3$ with $\det(T) = 1$ and no eigenvalue -1 then there is an orthonormal basis B of $\mathbb{R}^3$ for which

$$T_B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}.$$

Conclude T is a rotation by θ about some axis through the origin.

Solution. (i) Similar arguments to part (i) of Exercise 1.2.8, but using $ad-bc = 1$, imply $a = d$. Since $ac + bd = 0$ we now have $a(b+c) = 0$. If $a \neq 0$ then we immediately see $b = -c$, while if $a = 0$ the equations given by orthogonality of S become $b^2 = 1 = c^2$, while the determinant gives $-bc = 1$; since $b, c \neq 0$ (since $\det(T) \neq 0$) we see $-bc = c^2$, so $b = -c$.

(ii) Since -1 is not an eigenvalue of T it follows from Exercise 1.2.6 that 1 is an eigenvalue of T. Let e_1 be the (normalised) eigenvector corresponding to eigenvalue 1 and consider $T_{e_1^\perp}$. By Lemma 1.2.7(ii) we have $\det(T_{e_1^\perp}) = 1$, so part (i) of this exercise applies to $T_{e_1^\perp}$ for any orthonormal basis $\{e_2, e_3\}$ of $e_1^\perp$. Writing $a = \cos\theta$ and $b = \sin\theta$ for a suitable angle θ implies that T_B is of the required form, where $B = \{e_1, e_2, e_3\}$. It is now clear that T defines a rotation about the axis defined by e_1 and the origin through an angle θ. □

Now we have the result we were aiming for.

Proposition 1.2.10. *Let T be a linear operator on $\mathbb{R}^3$. The following are equivalent:*

(i) $T \in \mathrm{SO}(3, \mathbb{R})$;

(ii) T acts by rotation about some line through the origin.

Proof. (i) $\Longrightarrow$ (ii) By definition T is an orthogonal matrix on $\mathbb{R}^3$ and $\det(T) = 1$. If -1 is an eigenvalue of T then part (iii) of Exercise 1.2.8 shows T is a rotation about a line through the origin. If -1 is not an eigenvalue of T then part (ii) of Exercise 1.2.9 shows T is a rotation about a line through the origin.

(ii) $\Longrightarrow$ (i) Choose an orthonormal basis B of $\mathbb{R}^3$ such that

$$T_B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}.$$

Now calculate

$$T_B^t T_B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos^2\theta + \sin^2\theta & -\cos\theta\sin\theta + \cos\theta\sin\theta \\ 0 & -\sin\theta\cos\theta + \sin\theta\cos\theta & \sin^2\theta + \cos^2\theta \end{pmatrix} = I_n,$$

so T is orthogonal by Exercise 1.1.2. Finally,

$$\det(T) = 1\left(\cos^2\theta + \sin^2\theta\right) = 1,$$

so $T \in \mathrm{SO}(3, \mathbb{R})$. □

Exercise 1.2.11. *Prove the implication (ii) $\Longrightarrow$ (i) of Proposition 1.2.10.*

Chapter 2

First paradoxes

In this chapter we meet our first "paradoxes", all of which are based on duplicating some objects.

Exercise 2.0.1. *Look up a definition of the English word* paradox*. Do the examples below qualify as paradoxes? What about the Banach–Tarski paradox itself?*

Solution. According to the OED:

> *paradox*: a situation or statement that seems impossible or difficult to understand because it contains two opposite facts or characteristics.

Though the paradoxes below do not match with our intuition, they do not contain two opposing mathematical facts. □

2.1 Cardinality

For an infinite cardinal I it is known that $2I = I$. This clearly contradicts our intuition from familiar arithmetic, but does it qualify as a paradox?

Exercise 2.1.1. *Consider the unit ball in $\mathbb{R}^3$. Divide the ball in n pieces, where $n \geq 2$, in such a way that each piece has infinitely many points. Using the above fact about cardinal arithmetic to identify each point of a piece with two points to obtain a second copy of each piece. Reassemble the original pieces to form the original ball and the duplicates of each piece to form a duplicate of the original ball. The Banach–Tarski paradox is proved!*

What is wrong with this reasoning?

Solution. The vagueness surrounding n and the "duplication" are not at all satisfactory, but the main point is that the Banach–Tarski paradox allows only *rigid motions* to be used, and there is no reason for the duplication here to be given by rigid motions. □

2.2 Spokes on a wheel paradox

This "paradox" is explained by Weston [11]. The idea will appear again later when we are proving the Banach–Tarski paradox.

Let l denote the line $(0, 1)$ along the x-axis in $\mathbb{R}^2$, and let ρ denote anti-clockwise rotation about the origin in $\mathbb{R}^2$ by $1/10$ radians; each $\rho^n(l)$ $(n \geq 1)$ is then a radius of the unit circle at an angle of $n/10$ radians from the x-axis. The union $W := \sqcup_{n=1}^{\infty} \rho^n(l)$ looks like the collection of *spokes* on a bicycle wheel, except there are infinitely many of them (the square union symbol signifies that we are taking the union of a disjoint family of sets). However, we can make another spoke as follows: let ρ^{-1} denote a clockwise rotation by $1/10$ radians, so that

$$\rho^{-1}(W) = \bigsqcup_{n=0}^{\infty} \rho^n(l) = W \bigsqcup l.$$

So applying one rotation in the opposite direction added one more spoke to W.

Exercise 2.2.1. *Why did we choose ρ to be rotation by $1/10$ radians above?*

Solution. Any choice of angle will work as long as it is not a rational multiple of π. We are using that ρ^n is a rotation by $n/10$ radians, and $l \notin W$ since $n/10$ is different from 2π for every $n \in \mathbb{N}$. For example, if we chose ρ to be rotation by $\pi/2$ radians then $\rho^4(l) = l$, so W would contain only four elements and $\rho^{-1}(W) = W$. □

2.3 Paradoxical decomposition of the free group

Consider the free group on two generators $\mathbb{F}_2$, and write the generators as a and b. Though it is deceptively simple, what we do now foreshadows the Banach–Tarski paradox, and is in fact one of the main parts of the proof. We will divide $\mathbb{F}_2$ into five disjoint sets, and then use the action of $\mathbb{F}_2$ on itself to rearrange these disjoint sets in to two copies of $\mathbb{F}_2$.

Recall that elements of $\mathbb{F}_2$ are represented by reduced words on the generating set $\{a, b, a^{-1}, b^{-1}\}$. For each $c \in \{a, b, a^{-1}, b^{-1}\}$ we define

$$W_c := \{w \in \mathbb{F}_2 : w \text{ is a reduced word beginning on the left with } c\}.$$

Since every element of $\mathbb{F}_2$ except the identity word e belongs to exactly one of the sets W_c we may write

$$\mathbb{F}_2 = W_a \bigsqcup W_b \bigsqcup W_{a^{-1}} \bigsqcup W_{b^{-1}} \bigsqcup \{e\}. \tag{2.1}$$

Now we claim that

$$a^{-1}W_a = W_a \bigsqcup W_b \bigsqcup W_{b^{-1}} \bigsqcup \{e\}.$$

Indeed, let w be a reduced word which does not belong to $W_{a^{-1}}$, so that aw is a reduced word in W_a and $w = a^{-1}(aw) \in a^{-1}W_a$. For the other inclusion let w be a reduced word in $a^{-1}W_a$. If $w = e$ then we are done, otherwise $w = a^{-1}w_a$ for some reduced word $w_a = as_1 \cdots s_n$, where $s_1 \neq a^{-1}$ since w_a is a reduced word in W_a. It follows that $w = s_1 \cdots s_n$ is a reduced word in $W_a \sqcup W_b \sqcup W_{b^{-1}} \sqcup \{e\}$. Thus, by acting on W_a we have written $\mathbb{F}_2$ as a union using only two of the pieces in the union (2.1): $\mathbb{F}_2 = a^{-1}W_a \sqcup W_{a^{-1}}$. Arguing similarly with b in place of a we obtain second copy of $\mathbb{F}_2$ as a union using two other pieces in the union (2.1): $\mathbb{F}_2 = b^{-1}W_b \sqcup W_{b^{-1}}$.

Exercise 2.3.1. *The* Cayley graph *of a group G with generating set $S \cup S^{-1}$ is the graph which has a vertex for each element of G, and an edge joining the vertices g and h if and only if $h = gs$ for some $s \in S \cup S^{-1}$.*

(i) Draw the Cayley graph of $\mathbb{Z}$ (which is cyclic). What part of the graph corresponds to the set W from the spokes on a wheel paradox? What part corresponds to $\rho^{-1}(W)$?

(ii) Draw (or look up) the Cayley graph of $\mathbb{F}_2$ with $S = \{a, b\}$. Mark the parts of the graph corresponding to $W_a, W_b, W_{a^{-1}}$ and $W_{b^{-1}}$. Use this to visualise how we obtained two copies of $\mathbb{F}_2$ above.

Solution. (i) The Cayley graph of $\mathbb{Z}$ is below.

$$\leftarrow \cdots\cdots\cdot \;\text{——}\; -2 \;\text{——}\; -1 \;\text{——}\; 0 \;\text{——}\; \color{red}{1} \;\text{——}\; \color{red}{2} \;\text{——}\; \color{red}{\cdot}\cdots\cdots \rightarrow$$

The red vertices correspond to elements of W, while $\rho^{-1}(W)$ also contains the vertex labelled 0.

(ii) For the Cayley graph of $\mathbb{F}_2$ we link to wikipedia. Every vertex of the tree shape on the right side belongs to W_a, and we see that it can be identified with the three tree shapes which make up the upper, right and lower areas of the graph — the vertex labelled a and the vertex labelled e have the same connections above, to the right and to the left. Hence combining this with the left side tree gives a copy of $\mathbb{F}_2$. The upper tree can also be identified with the three trees which make up the left, upper and right areas of the graph — compare the connections above, left and right of the vertex labelled b and of the vertex labelled e. Hence combining this upper section of the tree with the lower tree gives a second copy of $\mathbb{F}_2$. $\square$

2.4 Paradoxical decompositions in general

What we have shown above is that free groups have a *paradoxical decomposition*, in the following sense.

Definition 2.4.1. *Let G be a group acting on a set X. Say that X is G-*paradoxical *if there are disjoint subsets $A_1, \ldots, A_n$ and $B_1, \ldots, B_m$ of X, elements $g_1, \ldots, g_n$ and $h_1, \ldots, h_m$ of G such that*

$$\bigcup_{i=1}^{n} g_i A_i = X = \bigcup_{j=1}^{m} h_j B_j.$$

Such a collection is called a paradoxical decomposition *of X. If $X = G$ and the action is by left multiplication then we say G is paradoxical.*

Theorem 2.4.2. *Suppose that G acts on a set X by $\cdot : G \times X \to X$. A* fixed point *of this action is $x \in X$ such that there is $g \in G$ with $g \cdot x = x$; we say the action has no non-trivial fixed points if $g \cdot x = x$ implies $g = e$. If G is paradoxical and the action has no non-trivial fixed points then X is G-paradoxical.*

Proof. Choose $M \subseteq X$ such that M contains exactly one element from each G-orbit. We show that $\{g \cdot M : g \in G\}$ is a partition of X. It is clear that $\cup_{g \in G} g \cdot M = X$: fix $x \in X$; since M contains a point from each G-orbit there is $g \in G$ with $g \cdot x \in M$, so $x \in g^{-1} \cdot M$. Now suppose $g, h \in G$ and $x, y \in M$ with $g \cdot x = h \cdot y$, so $(h^{-1}g) \cdot x = y$. Thus x and y belong to the same orbit, so by definition of M we must have $x = y$. Hence $(h^{-1}g) \cdot x = x$, so x is a fixed point; we assumed that the only fixed points of the action are trivial we must have $h^{-1}g = e$. We have shown that if $g \cdot M$ and $h \cdot M$ have non-empty intersection then $g = h$.

Take disjoint subsets $A_1, \dots, A_n$ and $B_1, \dots, B_n$ of X, elements $g_1, \dots, g_n$ and $h_1, \dots, h_m$ of G such that

$$\bigcup_{i=1}^{n} g_i A_i = G = \bigcup_{j=1}^{m} h_j B_j.$$

coming from paradoxicality of G, and define subsets of X by

$$A_i^X := \bigcup_{g \in A_i} g \cdot M \quad \text{and} \quad B_j^X := \bigcup_{h \in B_j} h \cdot M.$$

The sets A_i^X, B_j^X are pairwise disjoint (because of the claim above) and, using the properties of a paradoxical decomposition of G,

$$\bigcup_{i=1}^{n} g_i \cdot A_i^X = \bigcup_{i=1}^{n} g_i \cdot \left(\bigcup_{g \in A_i} g \cdot M \right) = \left(\bigcup_{i=1}^{n} g_i A_i \right) \cdot M = G \cdot M = X.$$

Similarly

$$\bigcup_{j=1}^{m} h_j \cdot B_j^X = \bigcup_{i=1}^{m} h_j \cdot \left(\bigcup_{g \in B_j} g \cdot M \right) = \left(\bigcup_{j=1}^{m} h_j B_j \right) \cdot M = G \cdot M = X.$$

□

Exercise 2.4.3. *Let G be a group acting on a set X with no non-trivial fixed points.*

(i) Suppose that the action of G on X is paradoxical. Show that G is paradoxical.

(ii) Suppose that H is a subgroup of G and $E \subset X$ is H-paradoxical. Show that E is G-paradoxical.

Hint. For (i) look at one of the G-orbits, and transfer the paradoxical decomposition of that orbit to G.

Solution. (i) Take subsets $A_1, \dots, A_n, B_1, \dots, B_m$ of X, elements $g_1, \dots, g_n$ and $h_1, \dots, h_m$ of G witnessing the paradoxical decomposition of X, fix $x_0 \in X$ and let $M := \{g \cdot x_0 : g \in G\}$ denote one of the orbits of this action. Define

$$A_i^G := \{g \in G : g \cdot x_0 \in A_i \cap M\}, \quad B_j^G := \{g \in G : g \cdot x_0 \in B_j \cap M\}.$$

These sets are pairwise disjoint: since the action has no non-trivial fixed points $s \cdot x_0 = t \cdot x_0$ if and only if $s = t$, so the above sets are pairwise disjoint because the A_i and B_j are. We have

$$g_i A_i^G = \{g \in G : g \cdot x_0 \in g_i \cdot A_i \cap M\},$$

and since

$$\bigcup_{i=1}^{n} g_i \cdot A_i \cap M = \left(\bigcup_{i=1}^{n} g_i \cdot A_i \right) \cap M = X \cap M = M$$

it follows that $\cup_{i=1}^{n} g_i A_i^G = G$. Similarly, we find that $\cup_{j=1}^{m} h_j B_j^G = G$.

(ii) Again take subsets $A_1, \ldots, A_n, B_1, \ldots, B_m$ of X, elements $g_1, \ldots, g_n$ and $h_1, \ldots, h_m$ of H witnessing the paradoxical action of H on X. Since $g_1, \ldots, g_n$ and $h_1, \ldots, h_m$ belong to G it is immediate that E is also G-paradoxical. $\square$

Chapter 3

The Banach–Tarski paradox

3.1 A free subgroup of $\mathrm{SO}(3, \mathbb{R})$

We already know that $\mathbb{F}_2$ has a paradoxical decomposition, so our idea is to look for an identification of $\mathbb{F}_2$ with rotations of the sphere. We follow Runde [5, Theorem 0.1.4].

Theorem 3.1.1. *There is a subgroup of* $\mathrm{SO}(3, \mathbb{R})$ *which is isomorphic to* $\mathbb{F}_2$*.*

Proof. Let θ be an anticlockwise rotation by $\cos^{-1}(\frac{1}{3})$ around the x-axis and ϕ an anticlockwise rotation by $\cos^{-1}(\frac{1}{3})$ around the z-axis; with respect to the standard orthonormal basis B of $\mathbb{R}^3$ these rotations (and their inverses) are given by

$$\theta_B^{\pm} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{3} & \mp\frac{2\sqrt{2}}{3} \\ 0 & \pm\frac{2\sqrt{2}}{3} & \frac{1}{3} \end{pmatrix} \quad \text{and} \quad \phi_B^{\pm} = \begin{pmatrix} \frac{1}{3} & \mp\frac{2\sqrt{2}}{3} & 0 \\ \pm\frac{2\sqrt{2}}{3} & \frac{1}{3} & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Clearly θ and ϕ belong to $\mathrm{SO}(3, \mathbb{R})$, so every (reduced) word on θ and ϕ also belongs to $\mathrm{SO}(3, \mathbb{R})$. Our task is to show that no reduced word on θ and ϕ acts as the identity on $\mathcal{S}^3$, since then the map $\mathbb{F}_2 \to \mathrm{SO}(3, \mathbb{R})$ given on generators by $a \mapsto \theta$ and $b \mapsto \phi$ extends to an injective group homomorphism.

Let w be a word on θ and ϕ which is not the empty word; we will show, by induction on the length of w, that there is a vector on which w never acts as the identity. First assume that w ends (on the right) with $\phi^{\pm}$. We claim that

$$w_B \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \frac{1}{3^k} \begin{pmatrix} a \\ b\sqrt{2} \\ c \end{pmatrix}, \tag{3.1}$$

where k is the length of w, $a, b, c \in \mathbb{Z}$ and $3 \nmid b$. Observe that this is sufficient to prove the result for such w. Suppose that $k = 1$, so $w = \phi^{\pm}$; then

$$w_B \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 1 \\ \pm 2\sqrt{2} \\ 0 \end{pmatrix},$$

as claimed. Now suppose that the claim holds for a word w' of length k, so $w = \theta^{\pm} w'$ or $w = \phi^{\pm} w'$ and w' satisfies (3.1) for integers a', b', c'. Calculations show that

$$w_B \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \frac{1}{3^{k+1}} \begin{pmatrix} a \\ b\sqrt{2} \\ c \end{pmatrix},$$

where a, b, c are given by

$$\begin{aligned} &a = a' \mp 4b',\ b = b' \pm 2a',\ c = 3c' &&\text{if } w = \phi^{\pm} w'; \\ &a = 3a',\ b = b' \mp 2c',\ c = c' \pm 4b' &&\text{if } w = \theta^{\pm} w'. \end{aligned}$$

It remains to check that $3 \nmid b$. This follows from $3 \nmid b'$, but some tedious case-checking is required: in each case apply $3 \nmid b'$ to what is obtained.

- if $w = \phi^{\pm}\theta^{\pm} v$ then $b = b' \mp 2a'$ with $3 \mid a'$;
- if $w = \theta^{\pm}\phi^{\pm} v$ then $b = b' \mp 2c'$ with $3 \mid c'$;
- if $w = \phi^{\pm}\phi^{\pm} v$ then $b = 2b' - 9b''$, where b'' is the integer from the form (3.1) of v;
- if $w = \theta^{\pm}\theta^{\pm} v$ then $b = 2b' - 9b''$, where b'' is the integer from the form (3.1) of v.

This completes the proof of the claim.

To finish we must also take care of the case when w ends (on the right) with $\theta^{\pm}$; but what we have shown above implies that such w never acts as the identity on $\phi_B \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$. □

Exercises 3.1.2. *(i) In the above result we chose rotations about perpendicular axes. Do you think that any pair of axes will work? Guess what property is required in order that rotations about these axes generate a free group.*

(ii) Draw a picture which explains why the rotation $\theta^{-1}\phi\theta$ *acts differently on* $\mathcal{S}^3$ *to the rotation* ϕ*.*

(iii) Explain why we chose the angle of rotation to be $\cos^{-1}(\frac{1}{3})$*.*

Solution. (i) It is clear that if the same axis is used for θ and ϕ then $\theta\phi = \phi\theta$. In the general case we refer to [10, Theorem 2.2], which gives a complete answer as follows. *Suppose* θ *and* ϕ *are rotations by the same angle* α*. The subgroup generated by these rotations is isomorphic to* $\mathbb{F}_2$ *if and only if either (a) the axes are perpendicular and* $\cos\alpha$ *is a rational different from* $0, \pm 1/2, \pm 1$*, or (b) the axes are distinct and the cosine of the angle between the axes is transcendental.*

(ii) The picture should indicate that a point ϕp on the sphere does not travel the same distance under θ^{-1} as the point p, because ϕp lies at a different latitude to p.

(iii) This number is not a rational multiple of π, which means that no words on θ (or ϕ) only have a fixed point (other than the x-axis, which must give fixed points). This is exactly the same reasoning as in Exercise 2.2.1. □

3.2 The Hausdorff paradox

We proved in Exercise 1.2.4 that $\mathrm{SO}(3, \mathbb{R})$ acts on the sphere $\mathcal{S}^3$, so the idea is to apply Theorem 2.4.2 to get a paradoxical decomposition of $\mathcal{S}^3$. Unfortunately Theorem 2.4.2 requires that the action has no fixed points.

Exercise 3.2.1. *(i) What are the fixed points of the action of* $\mathrm{SO}(3, \mathbb{R})$ *on* $\mathcal{S}^3$*?*

(ii) Explain how these fixed points cause a problem when we try to transfer the paradoxical decomposition of the subgroup of $\mathrm{SO}(3, \mathbb{R})$ *to* $\mathcal{S}^3$*.*

Hint. For (ii) you may want to look at Section 4 of [11].

Solution. (i) Let $T \in \mathrm{SO}(3, \mathbb{R})$, and observe that $x \in \mathcal{S}^3$ is fixed by T if and only if x lies on the axis about which T is a rotation. This is clear geometrically, and can also be seen by choosing an orthonormal basis $B = \{e_1, e_2, e_3\}$ of $\mathbb{R}^3$ such that

$$T_B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}.$$

In this case T_B is rotation about the axis defined by e_1, and clearly $T_B(\pm e_1) = \pm e_1$, so $\pm e_1 \in \mathcal{S}^3$ are fixed points. More generally,

$$T_B \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} a \\ b\cos\theta - c\sin\theta \\ b\sin\theta - c\cos\theta \end{pmatrix},$$

from which we see that the vector $(a,b,c)^t$ is a fixed point only if it is equal to $\pm e_1$.

(ii) Obviously Theorem 2.4.2 does not apply; our task is to see why it fails. Suppose $p \in \mathcal{S}^3$ is a fixed point of the action of $\mathrm{SO}(3,\mathbb{R})$, say $\rho(p) = p$ but ρ is not the identity. If it happens that p belongs to the set M we chose in the proof of Theorem 2.4.2 then $p \in M$ and $p \in \rho(M)$, so $\{g \cdot M : g \in G\}$ is no longer a partition of $\mathcal{S}^3$. $\square$

The following result was effectively discovered by Hausdorff [3], and is known as the *Hausdorff paradox*.

Theorem 3.2.2. *There is a countable set $D \subset \mathcal{S}^3$ such that $\mathcal{S}^3 \setminus D$ is $\mathrm{SO}(3,\mathbb{R})$-paradoxical.*

Proof. Let F be the subgroup of $\mathrm{SO}(3,\mathbb{R})$ which is isomorphic to $\mathbb{F}_2$, as found in Theorem 3.1.1, and let D denote the set of fixed points of the action of F on $\mathcal{S}^3$ (so by Exercise 3.2.1 D contains two points for each axis of rotation corresponding to an element of F). Since $\mathbb{F}_2$ is countable (by Exercise 1.1.7) it follows that D is countable. If we can prove that F acts on $\mathcal{S}^3 \setminus D$ with no non-trivial fixed points then we can apply Theorem 2.4.2 and Exercise 2.4.3.

First we must check that F does indeed act on $\mathrm{SO}(3,\mathbb{R}) \setminus D$ (*i.e.* that no point in this set is sent to D by the action). Suppose that $p \in \mathcal{S}^3$ and $\rho \in F$ are such that $\rho(p) \in D$, so by definition of D there is a $\psi \in F$, which is not the neutral element, such that $\psi(\rho(p)) = \rho(p)$. Hence $(\rho^{-1}\psi\rho)(p) = p$, and since it $\rho^{-1}\psi\rho \in F$ cannot be the neutral element we have $p \in D$. We have shown that if $p \in \mathcal{S}^3 \setminus D$ then $\rho(p) \in \mathcal{S}^3 \setminus D$ for all $\rho \in F$, so we have a well-defined action of F on $\mathcal{S}^3 \setminus D$. The action of F on $\mathrm{SO}(3,\mathbb{R}) \setminus D$ cannot have any fixed points, since all such fixed points lie in D.

We have shown that F is a subgroup of $\mathrm{SO}(3,\mathbb{R})$ which acts on $\mathcal{S}^3 \setminus D$ with no non-trivial fixed points. By Theorem 2.4.2 it follows that $\mathcal{S}^3 \setminus D$ is F-paradoxical, so by Exercise 2.4.3 $\mathcal{S}^3 \setminus D$ is $\mathrm{SO}(3,\mathbb{R})$-paradoxical. $\square$

Exercise 3.2.3. *We finished the proof of Theorem 3.2.2 by applying Theorem 2.4.2. Look at the proof of Theorem 2.4.2 and write down the sets $A_i, B_j \subset \mathcal{S}^3 \setminus D$ and the elements $g_i, h_j \in \mathrm{SO}(3,\mathbb{R})$ which satisfy Definition 2.4.1 in this case.*

Solution. Let M be a set containing an element from each orbit of the action of F on $\mathcal{S}^3 \setminus D$. Writing W_c for the set of words in F beginning on the left with c we have

$$\mathcal{S}^3 \setminus D = M \bigsqcup (W_\theta \cdot M) \bigsqcup (W_\phi \cdot M) \bigsqcup (W_{\theta^{-1}} \cdot M) \bigsqcup (W_{\phi^{-1}} \cdot M),$$

and we obtain two copies by writing

$$\mathcal{S}^3 \setminus D = (\theta^{-1} W_\theta \cdot M) \bigsqcup (W_{\theta^{-1}} \cdot M) \quad \text{and} \quad \mathcal{S}^3 \setminus D = (\phi^{-1} W_\phi \cdot M) \bigsqcup (W_{\phi^{-1}} \cdot M),$$

where $\theta, \phi \in \mathrm{SO}(3, \mathbb{R})$ are the rotations defined in Theorem 3.1.1. □

3.3 The Banach–Tarski paradox

The Banach–Tarski paradox almost follows immediately from the Hausdorff paradox, but there are some technicalities to take care of.

First of all we need to find a paradoxical decomposition of the whole of $\mathcal{S}^3$, not just of $\mathcal{S}^3 \setminus D$. The idea of this proof is the same as the one in the spokes on a wheel "paradox", just adapted to three dimensions.

Proposition 3.3.1. *The sphere $\mathcal{S}^3$ is $\mathrm{SO}(3, \mathbb{R})$-paradoxical.*

Proof. Choose a line ℓ through the origin in $\mathbb{R}^3$ which does not intersect the set D from Theorem 3.2.2. Since there are uncountably many lines through the origin and the set D is countable such a line ℓ certainly exists. We want to find an angle α_0 so that if σ is (anticlockwise) rotation about ℓ through an angle α_0 (note that $\sigma \in \mathrm{SO}(3, \mathbb{R})$) then the sets $\sigma^n(D)$ $(n \in \mathbb{N})$ are pairwise disjoint, like the sets $\rho^n(l)$ from the spokes on a wheel "paradox".

Let $\sigma_\alpha \in \mathrm{SO}(3, \mathbb{R})$ be anticlockwise rotation by angle α about the line ℓ, and consider

$$\{\alpha \in [0, 2\pi) : \text{ there is } p \in D \text{ and } n \in \mathbb{N} \text{ with } \sigma_\alpha^n(p) = \sigma_{n\alpha}(p) \in D\}.$$

For each pair $(p, q) \in D$ there is at most one angle $\alpha \in [0, 2\pi)$ with $\sigma_\alpha(p) = q$, so since $\mathbb{N}$ is countable each pair $(p, q) \in D \times D$ contributes only countably many elements to the set above. Since $D \times D$ is countable the above set is countable, so we may choose an angle α_0 which is not in the set. Let $\sigma := \sigma_{\alpha_0}$ denote the corresponding rotation, and observe that σ has the property we wanted: since $\sigma^n(D) \cap D$ is empty it follows that $\sigma^m(D) \cap \sigma^n(D)$ is empty for all $m, n \in \mathbb{N}$ with $m \neq n$.

To finish the proof we apply the same idea from the spokes on a wheel paradox. Let

$$E := \bigsqcup_{n=0}^{\infty} \sigma^n(D),$$

and note that $\mathcal{S}^3 \setminus D = (\mathcal{S}^3 \setminus \sigma^{-1}E) \sqcup E$; hence $\mathcal{S}^3 = (\mathcal{S}^3 \setminus \sigma^{-1}E) \sqcup \sigma^{-1}E$. Define pairwise disjoint subsets of $\mathcal{S}^3$ by

$$\begin{aligned}
&A_1 = (W_\theta \cdot M) \cap (\mathcal{S}^3 \setminus E) \cap \theta(\mathcal{S}^3 \setminus E), && B_1 = (W_\phi \cdot M) \cap (\mathcal{S}^3 \setminus E) \cap \phi(\mathcal{S}^3 \setminus E),\\
&A_2 = (W_\theta \cdot M) \cap E \cap \theta(\mathcal{S}^3 \setminus E), && B_2 = (W_\phi \cdot M) \cap E \cap \phi(\mathcal{S}^3 \setminus E),\\
&A_3 = (W_{\theta^{-1}} \cdot M) \cap (\mathcal{S}^3 \setminus E), && B_3 = (W_{\phi^{-1}} \cdot M) \cap (\mathcal{S}^3 \setminus E),\\
&A_4 = (W_\theta \cdot M) \cap (\mathcal{S}^3 \setminus E) \cap \theta(E), && B_4 = (W_\phi \cdot M) \cap (\mathcal{S}^3 \setminus E) \cap \phi(E),\\
&A_5 = (W_\theta \cdot M) \cap E \cap \theta(E), && B_5 = (W_\phi \cdot M) \cap E \cap \phi(E),\\
&A_6 = (W_{\theta^{-1}} \cdot M) \cap E, && B_6 = (W_{\phi^{-1}} \cdot M) \cap E.
\end{aligned}$$

Also define elements of $\mathrm{SO}(3,\mathbb{R})$:

$$\begin{aligned}
&g_1 = \theta^{-1}, g_2 = \theta^{-1}, g_3 = e, g_4 = \sigma^{-1}\theta^{-1}, g_5 = \sigma^{-1}\theta^{-1}, g_6 = \sigma^{-1}\\
&h_1 = \phi^{-1}, h_2 = \phi^{-1}, h_3 = e, h_4 = \sigma^{-1}\phi^{-1}, h_5 = \sigma^{-1}\phi^{-1}, h_6 = \sigma^{-1}.
\end{aligned}$$

In the following calculation we use that intersection distributes over union a number of times, *i.e.*

$$(X_i \cap Y_i) \cup (X_i \cap Z_i) = X_i \cap (Y_i \cup Z_i). \tag{3.2}$$

The sets involved on the left side are labelled in the calculation below using underbraces, and the resulting sets on the right side with overbraces. Now

$$(\theta^{-1}\cdot A_1)\bigsqcup(\theta^{-1}\cdot A_2)\bigsqcup A_3\bigsqcup(\sigma^{-1}\theta^{-1}\cdot A_4)\bigsqcup(\sigma^{-1}\theta^{-1}\cdot A_5)\bigsqcup(\sigma^{-1}\cdot A_6)$$

$$=\big(\theta^{-1}(A_1\sqcup A_2)\sqcup A_3\big)\bigsqcup\big(\sigma^{-1}\big(\theta^{-1}(A_4\sqcup A_5)\sqcup A_6\big)\big)$$

$$\begin{aligned}
=&\Bigg(\theta^{-1}\Big(\overbrace{\big(\underbrace{(W_\theta\cdot M)}_{X_1}\cap\underbrace{(\mathcal{S}^3\setminus E)\cap\theta(\mathcal{S}^3\setminus E)}_{Y_1}\big)}^{A_1}\sqcup\overbrace{\big(\underbrace{(W_\theta\cdot M)}_{X_1}\cap\underbrace{E\cap\theta(\mathcal{S}^3\setminus E)}_{Z_1}\big)}^{A_2}\Big)\\
&\qquad\qquad\qquad\qquad\sqcup\overbrace{\big((W_{\theta^{-1}}\cdot M)\cap(\mathcal{S}^3\setminus E)\big)}^{A_3}\Bigg)\\
&\bigsqcup\Bigg(\sigma^{-1}\Bigg(\theta^{-1}\Big(\overbrace{\big(\underbrace{(W_\theta\cdot M)}_{X_2}\cap\underbrace{(\mathcal{S}^3\setminus E)\cap\theta(E)}_{Y_2}\big)}^{A_4}\sqcup\overbrace{\big(\underbrace{(W_\theta\cdot M)}_{X_2}\cap\underbrace{E\cap\theta(E)}_{Z_2}\big)}^{A_5}\Big)\\
&\qquad\qquad\qquad\qquad\sqcup\overbrace{\big((W_{\theta^{-1}}\cdot M)\cap E\big)}^{A_6}\Bigg)\Bigg)\\
=&\Bigg(\theta^{-1}\Big(\overbrace{(W_\theta\cdot M)}^{X_1}\cap\Big(\overbrace{\big(\underbrace{(\mathcal{S}^3\setminus E)}_{Y_3}\cap\underbrace{\theta(\mathcal{S}^3\setminus E)}_{X_3}\big)}^{Y_1}\cup\overbrace{\big(\underbrace{E}_{Z_3}\cap\underbrace{\theta(\mathcal{S}^3\setminus E)}_{X_3}\big)}^{Z_1}\Big)\Big)\\
&\qquad\qquad\qquad\qquad\cup\big((W_{\theta^{-1}}\cdot M)\cap(\mathcal{S}^3\setminus E)\big)\Bigg)\\
&\bigsqcup\Bigg(\sigma^{-1}\Bigg(\theta^{-1}\Big(\overbrace{(W_\theta\cdot M)}^{X_2}\cap\Big(\overbrace{\big(\underbrace{(\mathcal{S}^3\setminus E)}_{Y_4}\cap\underbrace{\theta(E)}_{X_4}\big)}^{Y_2}\cup\overbrace{\big(\underbrace{E}_{Z_4}\cap\underbrace{\theta(E)}_{X_4}\big)}^{Z_2}\Big)\Big)\\
&\qquad\qquad\qquad\qquad\cup\big((W_{\theta^{-1}}\cdot M)\cap E\big)\Bigg)\Bigg)\\
=&\Bigg(\theta^{-1}\Big((W_\theta\cdot M)\cap\big(\overbrace{\theta(\mathcal{S}^3\setminus E)}^{X_3}\cap\big(\overbrace{(\mathcal{S}^3\setminus E)}^{Y_3}\cup\overbrace{E}^{Z_3}\big)\big)\Big)\\
&\qquad\qquad\qquad\qquad\cup\big((W_{\theta^{-1}}\cdot M)\cap(\mathcal{S}^3\setminus E)\big)\Bigg)\\
&\bigsqcup\Bigg(\sigma^{-1}\Bigg(\theta^{-1}\Big((W_\theta\cdot M)\cap\big(\overbrace{\theta(E)}^{X_4}\cap\big(\overbrace{(\mathcal{S}^3\setminus E)}^{Y_4}\cup\overbrace{E}^{Z_4}\big)\big)\Big)\\
&\qquad\qquad\qquad\qquad\cup\big((W_{\theta^{-1}}\cdot M)\cap E\big)\Bigg)\Bigg)
\end{aligned}$$

$$
\begin{aligned}
&= \big(\theta^{-1}\big((W_\theta \cdot M) \cap \theta(\mathcal{S}^3 \setminus E)\big) \cup \big((W_{\theta^{-1}} \cdot M) \cap (\mathcal{S}^3 \setminus E)\big)\big) \\
&\quad \bigsqcup \Big(\sigma^{-1}\Big(\theta^{-1}\Big((W_\theta \cdot M) \cap \theta(E)\Big) \cup \big((W_{\theta^{-1}} \cdot M) \cap E\big)\Big)\Big) \\
&= \left(\big(\underbrace{\theta^{-1}(W_\theta \cdot M)}_{Y_5} \cap \underbrace{(\mathcal{S}^3 \setminus E)}_{X_5}\big) \cup \big(\underbrace{(W_{\theta^{-1}} \cdot M)}_{Z_5} \cap \underbrace{(\mathcal{S}^3 \setminus E)}_{X_5}\big) \right) \\
&\quad \bigsqcup \left(\sigma^{-1} \left(\big(\underbrace{\theta^{-1}(W_\theta \cdot M)}_{Y_6} \cap \underbrace{E}_{X_6}\big) \cup \big(\underbrace{(W_{\theta^{-1}} \cdot M)}_{Z_6} \cap \underbrace{E}_{X_6}\big) \right) \right) \\
&= \left(\big(\overbrace{\theta^{-1}(W_\theta \cdot M)}^{Y_5} \cup \overbrace{(W_{\theta^{-1}} \cdot M)}^{Z_5}\big) \cap \overbrace{(\mathcal{S}^3 \setminus E)}^{X_5} \right) \\
&\quad \bigsqcup \left(\big(\overbrace{\theta^{-1}(W_\theta \cdot M)}^{Y_6} \cup \overbrace{(W_{\theta^{-1}} \cdot M)}^{Z_6}\big) \cap \overbrace{E}^{X_6} \right) \\
&= \big((\mathcal{S}^3 \setminus D) \cap (\mathcal{S}^3 \setminus E)\big) \bigsqcup \sigma^{-1}\big((\mathcal{S}^3 \setminus D) \cap E\big) \\
&= (\mathcal{S}^3 \setminus E) \bigsqcup \sigma^{-1}(E \setminus D) = (\mathcal{S}^3 \setminus E) \bigsqcup \sigma^{-1}(\sigma E) = \mathcal{S}^3.
\end{aligned}
$$

Repeating the above calculation shows $\sqcup_{j=1}^{6} h_j \cdot B_j = \mathcal{S}^3$. This completes the proof. □

Now we are ready to finish the proof of the Banach–Tarski paradox.

Exercise 3.3.2. *Can you guess how we will get a paradoxical decomposition of the solid ball from the paradoxical decomposition of $\mathcal{S}^3$ we found in Proposition 3.3.1? Explain the idea, along with any difficulties which you think may arise.*

Solution. We can get pairwise disjoint subsets of the solid ball from those of the sphere by extending inwards along each radius of the sphere: if A is one of the pieces in the paradoxical decomposition of $\mathcal{S}^3$ define $A^r := \{ta : t \in (0,1], a \in A\}$. The difficulty is the origin, which is not moved by any of the rotations we used for the sphere. □

We denote the solid unit ball in $\mathbb{R}^3$ centred at the origin by $\mathcal{B}^3 := \{x \in \mathbb{R}^3 : \|x\| \leq 1\}$. Write $\mathbb{E}^n$ for the n-dimensional Euclidean motion group, which consists of all translations and rotations.

Theorem 3.3.3. *The unit ball $\mathcal{B}^3$ in $\mathbb{R}^3$ is $\mathbb{E}^3$-paradoxical. That is, we can split the ball $\mathcal{B}^3$ in finitely many pieces, then rearrange these pieces using rotations and translations in $\mathbb{E}^3$ to get two copies of $\mathcal{B}^3$.*

Proof. Let $A_1, \dots, A_n$ and $B_1, \dots, B_m$ be pairwise disjoint subsets of $\mathcal{S}^3$ and $g_1, \dots, g_n, h_1, \dots, h_m \in \mathrm{SO}(3, \mathbb{R})$ giving the paradoxical decomposition of $\mathcal{S}^3$ found in Proposition 3.3.1. As in Exercise 3.3.2 define disjoint subsets A_i^r and B_j^r of $\mathcal{B}^3$ by filling in radially:

$$A_i^r := \{ta : t \in (0,1], a \in A_i\} \quad \text{and} \quad B_j^r := \{tb : t \in (0,1], b \in B_j\};$$

these sets, together with $g_1, \dots, g_n, h_1, \dots, h_m \in \mathrm{SO}(3, \mathbb{R})$ in above, give a paradoxical decomposition of $\mathcal{B}^3 \setminus \{0\}$.

It remains to fix the missing origin. Fortunately we can reuse the same idea from the spokes on a wheel paradox, in the same way as we did to fix the set D in Proposition 3.3.1: we will define a set C and a rotation τ such that $\mathcal{B}^3 = (\mathcal{B}^3 \setminus C) \sqcup C$, and $\mathcal{B}^3 \setminus \{0\} = (\mathcal{B}^3 \setminus C) \sqcup \tau C$. There are many choices for how to define C and τ — we just need to make sure that $C \subset \mathcal{B}^3$ and that rotation by τ has similar properties to the spokes on a wheel paradox. Let $x = (1/10, 0, 0)$ and choose a line through x which does not pass through the origin; this line will be the axis of rotation. Let τ denote clockwise rotation about this line through 1 radian. Note that τ is not an element of $\mathrm{SO}(3, \mathbb{R})$, but $\tau \in \mathbb{E}^3$, and $\tau^n(0) \neq 0$ for all $n \in \mathbb{N}$. For the same reason as we used in Section 2.2 and in Proposition 3.3.1, $C := \sqcup_{n=0}^{\infty} \tau^n(0)$ satisfies $\tau C = \sqcup_{n=1}^{\infty} \tau^n(0) = C \setminus \{0\}$. Combining this with the paradoxical decomposition of $\mathcal{B}^3 \setminus \{0\}$ gives a paradoxical decomposition of $\mathcal{B}^3$. One can write out the sets and group elements involved in this paradoxical decomposition formally, but it is extremely long! See Exercise 3.3.7. □

Exercise 3.3.4. *Why does the Banach–Tarski paradox not say that the unit ball in $\mathbb{R}^3$ is $\mathrm{SO}(3, \mathbb{R})$-paradoxical?*

Solution. The rotation we used to fix the missing origin was not a rotation about a line through the origin (the idea would not work on such a rotation, since they do not move the origin), hence this rotation was not an element of $\mathrm{SO}(3, \mathbb{R})$. □

Obviously the Banach–Tarski paradox as stated above does not require us to work with the *unit* ball in $\mathbb{R}^3$: the same idea by filling in the solid ball radially allows us to duplicate a solid ball of any radius.

You may have noticed that it quickly became difficult to keep track of the paradoxical decompositions of sets involved, so it is convenient to introduce some terminology (perhaps it would have been more convenient to do so before proving the Hausdorff paradox). This terminology also makes it easier to give a more general form of the Banach–Tarski paradox, in which we no longer work with solid balls.

Definition 3.3.5. *Let G be a group acting on a set X. We say that two subsets A and B of X are G-*equidecomposable*, and write $A \sim_G B$, if there are disjoint subsets $A_1, \dots, A_n$ and $B_1, \dots, B_n$ of X (we do not need any assumption on intersections $A_i \cap B_j$) and $g_1, \dots, g_n \in G$ such that $A = \sqcup_{i=1}^n A_i$ and $B = \sqcup_{j=1}^n B_j$ and $B_i = g_i \cdot A_i$ for $1 \le i \le n$. We also write $A \preceq_G B$ if A is G-equidecomposable with a subset of B.*

Lemma 3.3.6. *Let G be a group acting on a set X.*

(i) If $A, B \subset X$ with $A \sim_G B$ then there is a bijection $\beta : A \to B$ with $C \sim_G \beta(C)$ for all $C \subset A$.

(ii) If $A_1, A_2, B_1, B_2 \subset X$ with $A_1 \cap A_2 = \emptyset = B_1 \cap B_2$ such that $A_1 \sim_G B_1$ and $A_2 \sim_G B_2$ then $A_1 \cup A_2 \sim_G B_1 \cup B_2$.

Proof. (i) Let $A_1, \dots, A_n, B_1, \dots, B_n \subset X$ and $g_1, \dots, g_n \in G$ witness $A \sim_G B$. Given $C \subset A$ define

$$\beta : C \to B;\ \beta(c) := g_i \cdot c, \quad c \in A_i.$$

(ii) Let $C_1, \dots, C_n$ and $D_1, \dots, D_n$ be pairwise disjoint subsets of A_1 and B_1 respectively, and $g_1, \dots, g_n \in G$ such that $g_i \cdot C_i = D_i$; also let $E_1, \dots, E_m$ and $F_1, \dots, F_m$ be pairwise disjoint subsets of A_2 and B_2 respectively, and $h_1, \dots, h_m \in G$ such that $h_j \cdot E_j = F_j$. Since A_1 and A_2 are disjoint the sets C_i and E_j are together pairwise disjoint (intersections $C_i \cap E_j$ are empty), similarly for D_i and F_j, so it is clear that these sets together with g_i, h_j implement $A_1 \cup A_2 \sim_G B_1 \cup B_2$. □

Exercises 3.3.7. *Suppose that G is a group acting on the set X.*

(i) Show that G-equidecomposability is an equivalence relation on the collection of subsets of X.

(ii) Show that if $A \subset X$ is G-paradoxical and $A \sim_G B$ then B is G-paradoxical.

(iii) Show that the relation $\preceq_G$ is a reflexive and transitive relation on the equivalence classes of $\sim_G$.

(iv) (a) *Reformulate the definition of a G-paradoxical set in terms of equidecomposability.*

(b) *Reformulate Proposition 3.3.1 using* $\mathrm{SO}(3,\mathbb{R})$*-equidecomposability.*

(c) *Summarise the proof of the Banach–Tarski paradox using equidecomposability.*

Solution. (i) That $\sim_G$ is reflexive is trivial. For symmetry suppose $A \sim_G B$, with sets $A_1, \ldots, A_n$ and $B_1, \ldots, B_n$ and $g_1, \ldots, g_n \in G$ witnessing the equidecomposability. Since $g_i \cdot A_i = B_i$ it follows that $g_i^{-1} \cdot B_i = A_i$, so the sets $B_1, \ldots, B_n$ and $A_1, \ldots, A_n$ and elements $g_1^{-1}, \ldots, g_n^{-1} \in G$ witness $B \sim_G A$. Finally, suppose $A \sim_G B$, witnessed by $(A_i)_{i=1}^n, (B_i)_{i=1}^n$ and $(g_i)_{i=1}^n$, and that $B \sim_G C$, witnessed by $(B'_j)_{j=1}^m$, $(C_j)_{j=1}^m$ and $(h_j)_{j=1}^m$. For $1 \leq i \leq n$ and $1 \leq j \leq m$ define $D_{i,j} \subset B$ by $B_{i,j} := B_i \cap B'_j$. Now define

$$A_{i,j} := g_i^{-1} \cdot B_{i,j} \quad \text{and} \quad C_{i,j} := h_j \cdot B_{i,j}.$$

We show that the sets $A_{i,j}$ are pairwise disjoint. Suppose that $x \in A_{i,j} \cap A_{k,l}$, so in particular $x \in g_i^{-1} \cdot B_i \cap g_k^{-1} \cdot B_k = A_i \cap A_k$, so $i = k$ since the sets $A_1, \ldots, A_n$ are pairwise disjoint. It now follows that $g_i \cdot x \in B_{i,j} \cap B_{i,l}$, so in particular $g_i \cdot x \in B'_j \cap B'_l$, so $j = l$ since the sets $B'_1, \ldots, B'_m$ are pairwise disjoint. This proves $(i,j) = (k,l)$, so that the sets $A_{i,j}$ are pairwise disjoint. An almost identical argument shows that the sets $C_{i,j}$ are pairwise disjoint. We have

$$\begin{aligned}\bigcup_{i,j} A_{i,j} &= \bigcup_{i,j} g_i^{-1} \cdot D_{i,j} = \bigcup_{i,j} g_i^{-1} \cdot (B_i \cap B'_j) = \bigcup_i g_i^{-1} \cdot \Big(B_i \cap \big(\cup_j B'_j\big)\Big) \\ &= \bigcup_i g_i^{-1} \cdot (B_i \cap B) = \bigcup_i A_i = A,\end{aligned}$$

and similarly

$$\begin{aligned}\bigcup_{i,j} C_{i,j} &= \bigcup_{i,j} h_j \cdot D_{i,j} = \bigcup_{i,j} h_j \cdot (B_i \cap B'_j) = \bigcup_j h_j \cdot \Big(\big(\cup_i B_i\big) \cap B'_j\Big) \\ &= \bigcup_j h_j \cdot (B \cap B_j) = \bigcup_j C_j = C.\end{aligned}$$

Finally, it is clear that $(h_j g_i) \cdot A_{i,j} = C_{i,j}$, so these sets and elements of G witness that $A \sim_G C$.

(ii) Suppose that $A = \sqcup_{i=1}^n A_i$, $B = \sqcup_{i=1}^n B_i$, $g_i \cdot A_i = B_i$ (that is, $A \sim_G B$), and that A is G-paradoxical witnessed by pairwise disjoint subsets $C_1, \ldots, C_k$ and $D_1, \ldots, D_l$ of A and $s_1, \ldots, s_k, t_1, \ldots, t_l \in G$. Define subsets of B by

$$E_{i,j} := g_i \cdot \big((s_j^{-1} \cdot A_i) \cap C_j\big), \quad F_{i,j} := g_i \cdot \big((t_j^{-1} \cdot A_i) \cap D_j\big).$$

It is obvious that these sets are pairwise disjoint, from the pairwise disjointness of the sets involved. Define elements $p_{i,j} := g_i s_j g_i^{-1}, q_{i,j} := g_i t_j g_i^{-1} \in G$. We calculate

$$\begin{aligned}\bigcup_{i,j} p_{i,j} \cdot E_{i,j} &= \bigcup_{i,j} g_i s_j \big((s_j^{-1} \cdot A_i) \cap C_j\big) = \bigcup_i g_i \cdot \big(A_i \cap \big(\cup_j s_j \cdot C_j\big)\big) \\ &= \bigcup_i g_i \cdot \big(A_i \cap A\big) = \bigcup_i g_i \cdot A_i = B,\end{aligned}$$

and a similar calculation for $\cup_{i,j} q_{i,j} \cdot F_{i,j}$ shows that B is G-paradoxical.

(iii) It is obvious that $\preceq_G$ is reflexive. For transitivity, suppose $A \preceq_G B$ and $B \preceq_G C$. Take $A_1, \dots, A_n \subset A$ and $B_1^0, \dots, B_n^0 \subset B$, and elements $g_1, \dots, g_n \in G$, witnessing that $A \sim_G B^0$, for some $B^0 \subset B$. Also take $B_1, \dots, B_m \subset B$ and $C_1^0, \dots, C_m^0 \subset C$, and elements $h_1, \dots, h_m \in G$, witnessing that $B \sim_G C^0$, for some $C^0 \subset C$. Define $B_{i,j} := B_i^0 \cap B_j$ (note that the union of these sets is not necessarily all of B, only all of B^0), and define further

$$A_{i,j} := g_i^{-1} \cdot B_{i,j} \quad \text{and} \quad C_{i,j} := h_j \cdot B_{i,j}.$$

To see the above sets are pairwise disjoint suppose $x \in C_{i,j} \cap C_{k,l}$, so

$$x \in h_j \cdot B_{i,j} \cap h_l \cdot B_{k,l} \subset h_j \cdot B_j^0 \cap h_l \cdot B_l^0 \subset C_j^0 \cap C_l^0.$$

Since $C_1^0, \dots, C_m^0$ are pairwise disjoint it follows that $k = l$. Now we have $x \in C_{i,j} \cap C_{k,j}$, so

$$h_j^{-1} \cdot x \in B_{i,j} \cap B_{k,j} \subset B_i^0 \cap B_k^0,$$

which implies $i = k$. Hence $(i, j) = (k, l)$ so the sets $C_{i,j}$ are pairwise disjoint. Similar arguments show that the sets $A_{i,j}$ are pairwise disjoint. The union $\cup_{i,j} C_{i,j}$ is a subset of C, while

$$\begin{aligned}\bigcup_{i,j} A_{i,j} &= \bigcup_{i,j} g_i^{-1} \cdot (B_i^0 \cap B_j) = \bigcup_i g_i^{-1} \cdot \big(B_i^0 \cap \big(\cup_j B_j\big)\big) = \bigcup_i g_i^{-1} \cdot (B_i^0 \cap B) \\ &= \bigcup_i A_i = A.\end{aligned}$$

Since it is immediate that $(h_j g_i) \cdot A_{i,j} = C_{i,j}$ we have shown that $A \preceq_G C$.

(iv) A set is G-paradoxical if and only if it is G-equidecomposable with two copies of itself. Note that with this part (ii) becomes

$$B \sim_G A \sim_G A \sqcup A \sim_G B \sqcup B.$$

Proposition 3.3.1 shows that $\mathcal{S}^3 \setminus D$ is $\mathrm{SO}(3, \mathbb{R})$-equidecomposable with $\mathcal{S}^3$, *i.e.* $(\mathcal{S}^3 \setminus D) \sim_{\mathrm{SO}(3,\mathbb{R})} \mathcal{S}^3$.

Summarising Proposition 3.3.1:

$$\mathcal{S}^3 \sim_{\mathrm{SO}(3,\mathbb{R})} \left(\mathcal{S}^3 \setminus D\right) \sim_{\mathrm{SO}(3,\mathbb{R})} \left(\mathcal{S}^3 \setminus D\right) \bigsqcup \left(\mathcal{S}^3 \setminus D\right) \sim_{\mathrm{SO}(3,\mathbb{R})} \mathcal{S}^3 \bigsqcup \mathcal{S}^3.$$

Summarising the Banach–Tarski paradox, Theorem 3.3.3,

$$\mathcal{B}^3 \sim_{\mathbb{E}^3} \left(\mathcal{B}^3 \setminus \{0\}\right) \sim_{\mathbb{E}^3} \left(\mathcal{B}^3 \setminus \{0\}\right) \bigsqcup \left(\mathcal{B}^3 \setminus \{0\}\right) \sim_{\mathbb{E}^3} \mathcal{B}^3 \bigsqcup \mathcal{B}^3.$$

Note that the calculations in this part use Lemma 3.3.6 part (ii). □

Recall the Cantor–Schröder–Bernstein theorem for cardinals, which states that if $I \leq J$ and $J \leq I$ then $I = J$. The following theorem is a version of this result for the relations $\preceq_G$ and $\sim_G$.

Theorem 3.3.8. *Let G be a group acting on a set X. Suppose that A and B are subsets of X such that $A \preceq_G B$ and $B \preceq_G A$. Then $A \sim_G B$.*

Proof. Let $B' \subset B$ and $A' \subset A$ be such that $A \sim_G B'$ and $B \sim_G A'$. Let $\beta : A \to B'$ and $\gamma : B \to A'$ be bijections as in Lemma 3.3.6. Define $C_0 := A \setminus A'$, and define inductively $C_{n+1} := \gamma \circ \beta(C_n)$. Write $C := \cup_{n=0}^{\infty} C_n$. We have that $\gamma^{-1}(A \setminus C) = B \setminus \beta(C)$, which implies $(A \setminus C) \sim_G (B \setminus \beta(C))$. Similarly $C \sim_G \phi(C)$. Hence, by Lemma 3.3.6 again,

$$A = ((A \setminus C) \cup C) \sim_G ((B \setminus \beta(C)) \cup \beta(C)) = B.$$

□

Now we can give the strong form of the Banach–Tarski paradox. For a set $X \subset \mathbb{R}^n$, a point x is called an *interior point of* X if there is an open ball centred at x, say $B_\epsilon(x)$, with $B_\epsilon(x) \subset X$.

Theorem 3.3.9. *Any two bounded subsets of $\mathbb{R}^3$ with non-empty enterior are $\mathbb{E}^3$-equidecomposable.*

Proof. Let A and B be subsets of $\mathbb{R}^3$ with non-empty interior. We will show $A \preceq_{\mathbb{E}^3} B$; since A and B are arbitrary the same argument also shows $B \preceq_{\mathbb{E}^3} A$, so by Theorem 3.3.8 $A \sim_{\mathbb{E}^3} B$. Take solid balls of suitable radius K and L with $A \subset K$ and $L \subset B$ (this is possible because we assumed A was bounded and B has an interior point). Choose n large enough that K can be covered by n copies of L (the copies of L are allowed to have non-empty intersection). Let M denote a set of n disjoint copies of L then, by applying our first version of the Banach–Tarski paradox, Theorem 3.3.3, n times, $M \sim_{\mathbb{E}^3} L$. This means

$$A \subset K \preceq_{\mathbb{E}^3} M \preceq_{\mathbb{E}^3} L \subset B,$$

so by Exercise 3.3.7 part (iii) $A \preceq_{\mathbb{E}^3} B$. □

The Banach–Tarski paradox is closely related to the *axiom of choice*:

> **Axiom of choice**. For any collection of non-empty sets $\{X_i\}_{i \in I}$ there is a set X containing exactly one element from each X_i.

This proved controversial: Borel objected to the Hausdorff paradox because of its use of the axiom of choice, since the use of the axiom of choice means the use of a set which cannot be 'explicitly' defined. We will see later that the Banach–Tarski paradox necessarily involves sets which are not Lebesgue-measurable (this statement will be made precise in Chapter 4); it is now known that constructing a set which is not Lebesgue-measurable requires some form of the axiom of choice, so that the Banach–Tarski paradox also does require some form of the axiom of choice as an assumption. We refer to [10, Chapter 13] for a detailed account of the connection between the axiom of choice and the Banach–Tarski paradox, including references for the statements in this paragraph.

Exercises 3.3.10. *(i) What is your view on the axiom of choice?*

(ii) Identify the points in this project where the axiom of choice was used.

(iii) Look up the original paper of Banach and Tarski [1]. What was their view?

Solution. (i) ...

(ii) In Theorem 2.4.2 we used the axiom of choice to define the set M by choosing an element of each orbit. This set M can be uncomfortable to work with, since we know very little about its elements.

(iii) We refer to [10, Chapter 13] for the translation of the relevant part of [1]. They argued that, although the axiom of choice leads to the Banach–Tarski paradox, in their research they also needed the axiom of choice to prove the following intuitive result:

> Two different polygons, one contained in the other, are never equivalent by finite decomposition.

That is, they needed the axiom of choice to prove that $\mathbb{R}^2$ does not have a version of the Banach–Tarski paradox. They seem to be arguing that the axiom of choice should not be discounted because it leads to the non-intuitive Banach–Tarski paradox, since it also appears necessary to prove results which match fully with our intuition. Unfortunately the quoted result has since been shown to hold without using the axiom of choice, but perhaps a similar argument could be made based on the fact that the axiom of choice is required to make Lebesgue measure behave how we expect — without the axiom of choice Lebesgue measure may fail to be countably additive. References for this and other statements we made here are given in [10, Chapter 13]. □

Chapter 4

The problem of measure

Now we go back to measure theory and look at the original motivation for developing the Hausdorff and Banach–Tarski paradoxes. First we give basic definitions and introduce Lebesgue measure, then cover non-measurable sets, which leads us to consider the problem of measure.

4.1 Basic measure theory and Lebesgue measure

Definition 4.1.1. *Let X be a set. A σ-algebra on X is a collection $\mathcal{A}$ of subsets of X such that*

i. $X \in \mathcal{A}$*;*

ii. $\mathcal{A}$ *is closed under complements;*

iii. $\mathcal{A}$ *is closed under countable unions (and therefore also countable intersections).*

For any set X there are two obvious σ-algebras on X: the collection $\mathcal{P}(X)$ of all subsets of X and $\{\emptyset, X\}$. We now see how to construct other examples.

Exercise 4.1.2. *Let X be a set.*

(i) Show that the intersection of an arbitrary non-empty family of σ-algebras on X is a σ-algebra on X.

(ii) Let $\mathcal{F}$ be a family of subsets of X. Show that there is a smallest σ-algebra $\mathcal{A}_{\mathcal{F}}$ on X that contains $\mathcal{F}$, and explain why it is unique.

Solution. (i) Let $\mathcal{C}$ denote a non-empty collection of σ-algebras on X, and $\mathcal{A}$ the intersection of the σ-algebras in $\mathcal{C}$. Since X belongs to each σ-algebra in $\mathcal{C}$, it is clear that $X \in \mathcal{A}$. If $A \in \mathcal{A}$ then A belongs to each σ-algebra in $\mathcal{C}$, so the complement of A belongs to each σ-algebra in $\mathcal{C}$ and hence to $\mathcal{A}$. Finally, if $\{A_n\}_{n\in\mathbb{N}}$ a countable family of subsets of $\mathcal{A}$ then $\cup_{n\in\mathbb{N}} A_n$ belongs to every σ-algebra in $\mathcal{C}$, so $\cup_{n\in\mathbb{N}} A_n \in \mathcal{A}$.

(ii) By the smallest σ-algebra that contains $\mathcal{F}$ we mean that $\mathcal{A}_{\mathcal{F}}$ is a σ-algebra which contains $\mathcal{F}$, and such that if $\mathcal{A}$ is a σ-algebra on X containing $\mathcal{F}$ then $\mathcal{A}$ contains $\mathcal{A}_{\mathcal{F}}$.

Let $\mathcal{C}$ denote the collection of σ-algebras on X which contain $\mathcal{F}$; this collection is non-empty, since it contains $\mathcal{P}(X)$. By part (i) there is a smallest σ-algebra on X containing $\mathcal{F}$, and by definition $\mathcal{A}_{\mathcal{F}}$ is contained in every σ-algebra on X containing $\mathcal{F}$.

To see that $\mathcal{A}_{\mathcal{F}}$ is unique suppose there is another smallest σ-algebra $\mathcal{B}_{\mathcal{F}}$ on X. Then, according to the first paragraph above, $\mathcal{A}_{\mathcal{F}} \subset \mathcal{B}_{\mathcal{F}}$ and $\mathcal{B}_{\mathcal{F}} \subset \mathcal{A}_{\mathcal{F}}$, so $\mathcal{A}_{\mathcal{F}} = \mathcal{B}_{\mathcal{F}}$. □

Now we are able to introduce an important σ-algebra on Euclidean space. Recall that a set $X \subset \mathbb{R}^n$ is called *open* if for every $x \in X$ there is $\epsilon > 0$ such that $B_\epsilon(x) \subset X$, *i.e.* every point of X is an interior point of X.

Definition 4.1.3. *The σ-algebra on $\mathbb{R}^n$ generated by all open subsets of $\mathbb{R}^n$ is called the* Borel σ-algebra *on $\mathbb{R}^n$, denoted $\mathfrak{B}(\mathbb{R}^n)$. Equivalently, $\mathfrak{B}(\mathbb{R}^n)$ is the σ-algebra on $\mathbb{R}^n$ generated by all* (half-open) boxes *on $\mathbb{R}^n$ [2, Proposition 1.1.5], that is, generated by all sets in $\mathbb{R}^n$ of the form*

$$\{(x_1, \dots, x_n) \in \mathbb{R}^n : a_i < x_i \le b_i \text{ for } i = 1, \dots, n\}.$$

The collections of sets introduced above are the domain of the *measures* we now introduce.

Definition 4.1.4. *Let X be a set and $\mathcal{A}$ a σ-algebra on X. A* measure *on $(X, \mathcal{A})$ is a function*

$$\mu : \mathcal{A} \to [0, +\infty]$$

with the following properties:

i. $\mu(\emptyset) = 0$;

ii. μ *is* countably additive, *i.e. if $\{A_n\}_{n\in\mathbb{N}}$ is a countable collection of pairwise disjoint subsets of X which all belong to $\mathcal{A}$ then $\mu(\cup_{n\in\mathbb{N}} A_n) = \sum_{n\in\mathbb{N}} \mu(A_n)$.*

*If, in addition, G is a group acting on X then we say μ is G-*invariant *if*

$$\mu(g \cdot A) = \mu(A), \quad g \in G,\ A \in \mathcal{A}.$$

In the above situation the elements of $\mathcal{A}$ are called *measurable sets*, and those sets with measure 0 are called *null sets*. We often speak simply of a measure on X, omitting the σ-algebra when it is clear from context.

Examples 4.1.5. *i. For any set X there is a measure on $\mathcal{P}(X)$ called* counting measure*, and defined by*

$$\mu(A) := \begin{cases} |A| & \text{if } A \text{ is finite;} \\ +\infty & \text{if } A \text{ is infinite.} \end{cases}$$

ii. Every probability space corresponds to a set X equipped with a σ-algebra and a measure μ satisfying $\mu(X) = 1$. In this case the measurable sets represent events, and the measure gives the probability of an event occurring.

iii. Every locally compact group carries a natural measure on the σ-algebra generated by all open sets, called Haar measure *(see Appendix B).*

Now we give a brief explanation of Lebesgue measure. Let us call a subset B of $\mathbb{R}^n$ a *box* if one can write

$$B = \{(x_1, \ldots, x_n) : x_i \in I_i\} = I_1 \times \cdots \times I_n,$$

where each $I_i \subset \mathbb{R}$ is an interval (we do not worry about whether the I_i are open or closed or half-open). It is natural that if $I = (a, b)$ (or $[a, b]$ or $(a, b]$ or $[a, b)$) with $a \leq b$ then the length of I should be $\mathrm{len}(I) = b - a$. Similarly, we define the volume of a box $B = I_1 \times \cdots \times I_n$ in $\mathbb{R}^n$ to be

$$\mathrm{vol}(B) := \prod_{i=1}^{n} \mathrm{len}(I_i).$$

If a set $A \subset \mathbb{R}^n$ is contained in the union of countably many boxes $\{B_m\}_{m \in \mathbb{N}}$ then, according to the properties of measures, we should have $\mathrm{vol}(A) \leq \sum_{m \in \mathbb{N}} \mathrm{vol}(B_m)$. Now we can define *Lebesgue outer measure* on $\mathbb{R}^n$, denoted λ^*, by

$$\lambda^*(A) := \inf \left\{ \sum_{m \in \mathbb{N}} \mathrm{vol}(B_m) : B_m \text{ are boxes and } A \subset \cup_{m \in \mathbb{N}} B_m \right\}.$$

The following result summarises the results on Lebesgue measure in [2, Section 1.4] and defines Lebesgue measure.

Theorem 4.1.6. *There is a measure λ on $(\mathbb{R}^n, \mathfrak{B}(\mathbb{R}^n))$, which assigns to each box B its volume. This measure, which we call* Lebesgue measure*, is the unique measure μ on $(\mathbb{R}^n, \mathfrak{B}(\mathbb{R}^n))$ for which $\mu(B) = \mathrm{vol}(B)$ for all boxes B. Moreover, Lebesgue measure is translation-invariant, and agrees with Lebesgue outer measure where the latter is defined.*

We make two comments here on the properties of Lebesgue measure which we will not need later. First, it is possible to define a σ-algebra on $\mathbb{R}^n$ which contains $\mathfrak{B}(\mathbb{R}^n)$ on which Lebesgue measure is also defined; this σ-algebra and the resulting Lebesgue measure are called the completions of $\mathfrak{B}(\mathbb{R}^n)$ and λ, respectively. The latter is also called Lebesgue measure; it fixes the problem that there may be subsets of elements of $\mathfrak{B}(\mathbb{R}^n)$ which are not in $\mathfrak{B}(\mathbb{R}^n)$. Secondly, we stated that Lebesgue measure on $\mathfrak{B}(\mathbb{R}^n)$ is unique; in fact, if μ is any non-zero measure on $\mathfrak{B}(\mathbb{R}^n)$ that is finite on bounded sets and translation-invariant then μ is called a *Haar measure on* $\mathbb{R}^n$, and there is $c > 0$ for which $\mu = c\lambda$. In other words, Lebesgue measure is a Haar measure on $(\mathbb{R}^n, \mathfrak{B}(\mathbb{R}^n))$ (see Appendix B) and the requirement that it gives the unit cube a volume 1 determines the constant c.

Exercise 4.1.7. *Theorem 4.1.6 states that Lebesgue measure is translation-invariant.*

(i) Explain the statement: Lebesgue measure is invariant under the natural action of the group $\mathbb{R}^n$ on $\mathbb{R}^n$, defined in Exercise 1.2.2.

(ii) Show that Lebesgue measure is invariant under the action of the Euclidean motion group $\mathbb{E}^n$.

Solution. (i) The natural action defined in Exercise 1.2.2 is the translation action in the case of $\mathbb{R}^n$.

(ii) The Euclidean motion group is the group consisting of rotations and translations. Since every element of this group can be written as a combination of translations (for which we know Lebesgue measure is invariant) and rotations which fix the origin, it suffices to check that $\lambda(S(E + x)) = \lambda(E)$ for each Lebesgue measurable set E, $x \in \mathbb{R}^n$ (acting as translation) and a rotation S which fixes the origin. Using translation-invariance of λ we have, for $E \in \mathfrak{B}(n)$,

$$\lambda\big(S(E + x)\big) = \lambda\big(S(E) + S(x)\big) = \lambda\big(S(E)\big).$$

This shows that $\lambda \circ S$ is also a translation-invariant measure on the Borel sets, so by Theorem 4.1.6 $\lambda \circ S = c\lambda$ for some $c > 0$. Since S fixes the origin we have $S(\mathcal{B}^n) = \mathcal{B}^n$, so $c = 1$. $\square$

4.2 Vitali sets and the problem of measure

Is it possible that all subsets of $\mathbb{R}^n$ are Lebesgue measurable? This natural question was an open problem until it was solved by Vitali [8]. Make sure not to confuse "not Lebesgue measurable" with "has measure zero".

Theorem 4.2.1. *There is a subset of* $\mathbb{R}$ *which is not Lebesgue measurable.*

Proof. Define a relation $\sim$ on $\mathbb{R}$ by

$$x \sim y \iff x - y \in \mathbb{Q}.$$

First we check that this is an equivalence relation. Reflexivity is clear since $0 \in \mathbb{Q}$. If $x - y \in \mathbb{Q}$ then $y - x = -(x - y) \in \mathbb{Q}$, so $\sim$ is symmetric. For transitivity suppose $x \sim y$ and $y \sim z$, so $(x - y), (y, z) \in \mathbb{Q}$; thus $x - z = (x - y) + (y - z) \in \mathbb{Q}$ as $\mathbb{Q}$ is closed under addition. Each equivalence class of $\sim$ is of the form $\mathbb{Q} + x$ for some $x \in \mathbb{R}$, so each equivalence class is dense in $\mathbb{R}$. It also follows that each equivalence class intersects the interval $(0, 1)$, so since they are disjoint we may use the axiom of choice to form a set $V \subset (0, 1)$ containing exactly one element of each equivalence class. Now we prove this set V is not Lebesgue measurable.

Since $\mathbb{Q}$ is countable we can enumerate the set $\mathbb{Q} \cap (-1, 1)$, say by $\{r_n\}_{n\in\mathbb{N}}$, and define $V_n := V + r_n$. First we show that the sets $\{V_n\}_{n\in\mathbb{N}}$ are pairwise disjoint. If $V_m \cap V_n$ is not empty then there are $v, w \in V$ with $v + r_m = w + r_n$, so $v \sim w$ and therefore $m = n$, since distinct equivalence classes are always disjoint. Secondly, observe that

$$\bigcup_{n\in\mathbb{N}} V_n \subset (-1, 2), \tag{4.1}$$

since $V \subset (0, 1)$ and $-1 < r_n < 1$ for each $n \in \mathbb{N}$. Finally we show $(0, 1) \subset \cup_{n\in\mathbb{N}} V_n$. Let $x \in (0, 1)$ and take $v \in V$ so that $x \sim v$; thus $x - v \in \mathbb{Q}$ and $-1 < x - v < 1$, so $x - v = r_n$ for some $n \in \mathbb{N}$. Hence $x \in V_n$.

Now suppose that V is Lebesgue measurable. Since Lebesgue measure is translation-invariant $\lambda(V_n) = \lambda(V)$, and the disjointness of the sets $\{V_n\}_{n\in\mathbb{N}}$ implies

$$\lambda\left(\bigcup_{n\in\mathbb{N}} V_n\right) = \sum_{n\in\mathbb{N}} \lambda(V_n) = \sum_{n\in\mathbb{N}} \lambda(V). \tag{4.2}$$

If $\lambda(V) = 0$ then we have $\lambda(\cup_{n\in\mathbb{N}} V_n) = 0$, contradicting the above fact that $(0, 1) \subset \cup_{n\in\mathbb{N}} V_n$. If $\lambda(V) \neq 0$ then equation (4.2) implies $\lambda(V) = +\infty$, contradicting (4.1). The set V in this proof is called a *Vitali set.* $\square$

Exercise 4.2.2. *Give a construction of a Vitali set in* $\mathbb{R}^n$*.*

Solution. Define the relation $\overset{n}{\sim}$ on $\mathbb{R}^n$ by

$$(x_1, \ldots, x_n) \overset{n}{\sim} (y_1, \ldots, y_n) \iff x_i - y_i \in \mathbb{Q} \text{ for each } 1 \leq i \leq n.$$

This is an equivalence relation, and each equivalence class is of the form $\mathbb{Q}^n + x$ for some $x \in \mathbb{R}^n$. Form the set V by choosing one element of each equivalence class which also belongs to $(0,1)^n$. Enumerate the countable set $\mathbb{Q}^n \cap (-1,1)^n$, say by $\{r_m\}_{m\in\mathbb{N}}$. Similar arguments show that the sets $\{V + r_m\}_{m\in\mathbb{N}}$ are pairwise disjoint, and $(0,1)^n \subset \sup_{m\in\mathbb{N}}(V + r_m) \subset (-1,2)^n$. Then one produces a contradiction using the properties of measures in the same way. □

At the conclusion of proof of Theorem 4.2.1 it was essential that Lebesgue measure is *countably additive*: we used countable additivity to produce a contradiction in the case $\lambda(V) \neq 0$. After Vitali's result appeared in 1905 mathematicians were led to ask about what would happen if we removed the requirement of *countable* additivity.

Definition 4.2.3. *Let X be a set. An* algebra *on X is a collection $\mathcal{A}$ of subsets on X satisfying conditions (i),(ii) and (iii) in Definition 4.1.1, except that we require only finite unions in (iii). That is, an* algebra on X *is a collection of subsets of X containing X, closed under complements and finite unions. A* finitely additive measure on $(X, \mathcal{A})$ *is a function*

$$\mu : \mathcal{A} \to [0, +\infty]$$

with the following properties:

i. $\mu(\emptyset) = 0$*;*

ii. μ *is* (finitely) additive*, i.e. if $\{A_k\}_{k=1}^n$ is a collection of pairwise disjoint subsets of X which all belong to $\mathcal{A}$ then $\mu(\cup_{k=1}^n A_k) = \sum_{k=1}^n \mu(A_k)$.*

Theorem 4.2.1 leads us to consider the following *problem of measure.*

> **Problem of measure.** Is there a finitely additive measure on $\mathbb{R}^n$ which is invariant under the action of the group of isometries of $\mathbb{R}^n$, assigns the unit cube $[0,1]^n$ the measure 1, and which is defined on every subset of $\mathbb{R}^n$?

Sometimes the problem of measure is phrased as asking if Lebesgue measure can be extended to a finitely additive, isometry-invariant measure on all of $\mathbb{R}^n$ which normalises the unit cube.

In fact, we have already solved the problem of measure in $\mathbb{R}^3$.

Until now it appeared that the Banach–Tarski paradox creates a contradiction with Lebesgue measure: the unit ball in $\mathbb{R}^3$ has volume $\frac{4}{3}\pi$, while two copies of the unit ball have volume $\frac{8}{3}\pi$, yet we passed from one copy to two copies using only translations and rotations. But the Lebesgue measure of a set is supposed to be invariant under the action of the group containing all such translations and rotations. The following exercise resolves this confusion.

Exercise 4.2.4. *(i) Using the example of Vitali sets, explain why (at least) one of the sets involved in the Banach–Tarski paradox is not measurable for any finitely additive measure satisfying the conditions in the problem of measure.*

(ii) Explain why the Banach–Tarski paradox answers the problem of measure for $\mathbb{R}^3$.

Solution. (i) For any collection $\{C_i\}_{i=1}^m$ of disjoint measurable subsets of $\mathcal{B}^n$ and any elements $g_1, \ldots, g_m \in \mathbb{E}^n$ we have

$$\mu(\mathcal{B}^n) \geq \mu\big(\sqcup_{i=1}^m C_i \big) = \sum_{i=1}^m \mu(C_i) = \sum_{i=1}^m \mu(g_i \cdot C_i)$$

for any finitely additive and $\mathbb{E}^n$-invariant measure μ. Now let $A_1, \ldots, A_n$, $B_1, \ldots, B_m \subset \mathcal{B}^3$ and $g_1, \ldots, g_n, h_1, \ldots, h_m \in \mathbb{E}^3$ witness that $\mathcal{B}^3$ is $\mathbb{E}^3$-paradoxical. Then

$$\begin{aligned}
\mu(\mathcal{B}^3) &\geq \mu\left((\sqcup_{i=1}^n A_i) \bigsqcup (\sqcup_{j=1}^m B_j) \right) = \sum_{i=1}^n \mu(A_i) + \sum_{j=1}^m \mu(B_j) \\
&= \sum_{i=1}^n \mu(g_i \cdot A_i) + \sum_{j=1}^m \mu(h_j \cdot B_j) \\
&\geq \mu\left((\cup_{i=1}^n g_i \cdot A_i) \bigsqcup (\cup_{j=1}^m h_j \cdot B_j) \right) = 2\mu(\mathcal{B}^3)
\end{aligned}$$

for any measure μ which is $\mathbb{E}^3$-invariant. This can only happen if $\mu(\mathcal{B}^3) = 0$ or $\mu(\mathcal{B}^3) = \infty$, which cannot happen if μ also normalises the unit cube. It follows that at least one of the sets involved in the paradoxical decomposition above is not μ-measurable.

(ii) By (i) the Banach–Tarski paradox shows that for any measure μ satisfying the conditions of the problem of measure there must be a set in $\mathbb{R}^3$ which is not μ-measurable. Therefore the problem of measure has a negative answer in $\mathbb{R}^3$. $\square$

4.3 Tarski's theorem

Exercise 4.2.4 effectively shows that paradoxical decompositions prevent the existence of non-trivial finitely additive invariant measures defined on all subsets. Tarski [7] proved the converse to this result, which shows that the only obstruction to the existence of such measures is paradoxical decompositions. We do not give the proof of the hard direction, since it requires too much extra background.

Theorem 4.3.1. *Let G be a group acting on a set X and let $E \subset X$. The following are equivalent:*

(i) there is a finitely additive G-invariant measure on X which gives E the measure 1;

(ii) E is not G-paradoxical.

Proof. (i) $\Longrightarrow$ (ii) We show that if E is paradoxical then such measure cannot exist. Let $A_1, \dots, A_n, B_1, \dots, B_m \subset E$ and $g_1, \dots, g_n, h_1, \dots, h_m \in G$ witness that E is G-paradoxical. Then

$$\begin{aligned}\mu(E) &\geq \mu\left((\sqcup_{i=1}^n A_i) \bigsqcup (\sqcup_{j=1}^m B_j)\right) = \sum_{i=1}^n \mu(A_i) + \sum_{j=1}^m \mu(B_j) \\ &= \sum_{i=1}^n \mu(g_i \cdot A_i) + \sum_{j=1}^m \mu(h_j \cdot B_j) \\ &\geq \mu\left((\cup_{i=1}^n g_i \cdot A_i) \bigsqcup (\cup_{j=1}^m h_j \cdot B_j)\right) = 2\mu(E)\end{aligned}$$

for any measure μ which is G-invariant. This implies that $\mu(E)$ is 0 or $+\infty$.

(ii) $\Longrightarrow$ (i) Omitted. See [10, Corollary 9.2]. $\square$

Chapter 5

Amenable groups

In this section we will use the term *discrete group* to indicate a group with the discrete topology.

You will notice that one of the most important steps in our proof of the Banch–Tarski paradox was finding a free subgroup of the rotation group $\mathrm{SO}(3, \mathbb{R})$ in Theorem 3.1.1. Similarly, in Exercise 2.4.3 we saw that if a set is G-paradoxical then G is itself paradoxical. Therefore we may view the existence of paradoxical decompositions as being a statement about how complicated the group involved is. This was recognised by von Neumann, who abstracted this property in the following important definition.

Definition 5.0.1. *Let G be a discrete group. Say G is* amenable *is there is a finitely additive measure μ on all subsets of G which is G-invariant and normalises G; i.e. $\mu(gE) = \mu(E)$ for all $E \subset G$ and $\mu(G) = 1$.*

The idea of this definition is summarised in the following exercise.

Exercise 5.0.2. *Show that a discrete group is amenable if and only if it is not paradoxical.*

Solution. Let G be paradoxical, and suppose μ is a G-invariant measure defined on all subsets of G. If $A_1, \ldots, A_n, B_1, \ldots, B_m \subset G$ and elements $g_1, \ldots, g_n, h_1, \ldots, h_m \in G$ witness paradoxicality of G then

$$\begin{aligned}\mu(G) \geq \mu\left((\sqcup_{i=1}^n A_i) \sqcup (\sqcup_{j=1}^m B_j)\right) &= \sum_{i=1}^n \mu(A_i) + \sum_{j=1}^m \mu(B_j)\\ &= \sum_{i=1}^n \mu(g_i A_i) + \sum_{j=1}^m \mu(h_j B_j) = 2\mu(G),\end{aligned}$$

so $\mu(G) \neq 1$. Similar arguments prove the converse. □

In particular, we have the following fundamental example.

Example 5.0.3. *The free group $\mathbb{F}_n$ is not amenable when $n \geq 2$.*

For a space X with counting measure we define

$$\ell^\infty(X) := \left\{ \phi : X \to \mathbb{C} : \sup_{x \in X} |\phi(x)| < \infty \right\},$$

which is an *algebra* with pointwise operations, that is, it is a vector space (infinite-dimensional when X is infinite) and a ring in a compatible way. The vector space and ring operations are called the *pointwise operations*:

$$(c\phi)(x) := c(\phi(x)), \quad (\phi + \psi)(x) := \phi(x) + \psi(x), \quad (\phi\psi)(x) := \phi(x)\psi(x).$$

Exercise 5.0.4. *Suppose that G is a group, X a set and $\cdot : G \times X \to X$ an action of G on X. Show that*

$$\cdot : G \times \ell^\infty(X) \to \ell^\infty(X);\ (g \cdot \phi)(x) := \phi(g^{-1} \cdot x), \quad g \in G,\ x \in X,\ \phi \in \ell^\infty(X),$$

defines an action of G on $\ell^\infty(X)$ which is linear and multiplicative, that is, for each $g \in G$ the map $\phi \mapsto g \cdot \phi$ is linear and $g \cdot (\phi\psi) = (g \cdot \phi)(g \cdot \psi)$.

Solution. First we check that this is a group action:

$$(e \cdot \phi)(x) = \phi(e^{-1} \cdot x) = \phi(e \cdot x) = \phi(x),$$

and

$$\big(g \cdot (h \cdot \phi)\big)(x) = (h \cdot \phi)(g^{-1} \cdot x) = \phi\big((h^{-1}g^{-1}) \cdot x\big) = \phi\big((gh)^{-1} \cdot x\big) = (gh) \cdot x,$$

so we do have a group action if we can show that $g \cdot \phi$ satisfies the supremum condition when ϕ does. Suppose $\sup_{x \in X} |\phi(x)| < \infty$, so

$$\sup_{x \in X} \big|(g \cdot \phi)(x)\big| = \sup_{x \in X} |\phi(g^{-1} \cdot x)| = \sup_{y \in X} |\phi(y)| < \infty,$$

as required. Now, for $\phi, \psi \in \ell^\infty(X)$, $c \in \mathbb{C}$ and $g \in G$ we have

$$\begin{aligned} \big(g \cdot (c\phi + \psi)\big)(x) &= (c\phi + \psi)(g^{-1} \cdot x) = (c\phi)(g^{-1} \cdot x) + \psi(g^{-1} \cdot x) \\ &= c\big(g \cdot \phi\big)(x) + (g \cdot \psi)(x), \end{aligned}$$

so the action is linear, and

$$g \cdot (\phi\psi)(x) = (\phi\psi)(g^{-1} \cdot x) = \phi(g^{-1} \cdot x)\psi(g^{-1} \cdot x) = \big((g \cdot \phi)(g \cdot \psi)\big)(x).$$

□

Amenable groups are normally defined in a different way; the following exercise shows that our definition is the same as the more common one.

Exercise 5.0.5. *Let G be a discrete group. Show that the following conditions are equivalent:*

(i) G is amenable;

(ii) there is a map $I : \ell^\infty(G) \to \mathbb{C}$ which is:

(a) linear,

(b) contractive ($|I(\phi)| \leq \sup_{g\in G} |\phi(g)|$ for all $\phi \in \ell^\infty(G)$),

(c) positive (if $\phi(g) \geq 0$ for all $g \in G$ then $I(\phi) \geq 0$),

(d) G-invariant (for each $r \in G$ and $\phi \in \ell^\infty(G)$ we have $I(r\cdot\phi) = I(\phi)$).

Solution. (i) $\Longrightarrow$ (ii) Let μ denote the measure on G arising in (i). Define

$$I_\mu : \ell^\infty(G) \to \mathbb{C} : I_\mu(\phi) := \int_G \phi(g)\, d\mu(g), \quad \phi \in \ell^\infty(G),$$

that is, I_μ is integration against the measure μ, as defined before. We show that I_μ is the required map. That I_μ is linear and positive is a general fact about this integral. For contractivity take $\phi \in \ell^\infty(G)$ and calculate

$$\begin{aligned}|I_\mu(\phi)| &= \left|\int_G \phi(g)\, d\mu(g)\right| \leq \int_G \big|\phi(g)\big|\, d\mu(g) \leq \int_G \sup_{h\in G} \big|\phi(h)\big|\, d\mu(g) \\ &= \sup_{h\in G} \big|\phi(h)\big| \mu(G) = \sup_{h\in G} \big|\phi(h)\big|.\end{aligned}$$

Finally, invariance follows from invariance of μ:

$$\begin{aligned}I_\mu(r\cdot\phi) &= \int_G (r\cdot\phi)(g)\, d\mu(g) = \int_G \phi(r^{-1}g)\, d\mu(g) = \int_G \phi(h)\, d\mu(rh) \\ &= \int_G \phi(h)\, d\mu(h) = I_\mu(\phi).\end{aligned}$$

In the above calculation we changed variable $h = rg$, and replaced $d\mu(rh)$ by $d\mu(h)$ because $\mu(rA) = \mu(A)$ for each $A \subset G$.

(ii) $\Longrightarrow$ (i) Given such a map I, define a measure μ_I on X by

$$\mu_I(A) := I(\chi_A), \quad A \subset X.$$

We check that μ_I satisfies the conditions in Definition 5.0.1. Clearly μ_I is defined on all subsets of G. For finite additivity note that if $A, B \subset G$ are disjoint then $\chi_{A\cup B} = \chi_A + \chi_B$, so

$$\mu_I(A \cup B) = I(\chi_{A\cup B}) = I(\chi_A + \chi_B) = I(\chi_A) + I(\chi_B) = \mu_I(A) + \mu_I(B);$$

finite additivity follows. That μ_I is a finitely additive measure follows because $I(\chi_\emptyset) = I(0) = 0$ and I is positive, so $\mu_I(A) \geq 0$ for all $A \subset G$. Also,

$$\mu_I(G) = I(\chi_G) = 1,$$

the final equality following because I is contractive. Finally, for $r \in G$ and $A \subset G$,

$$\mu_I(rA) = I(\chi_{rA}) = I(r \cdot \chi_A) = I(\chi_A) = \mu_I(A),$$

since I is G-invariant. (In the second equality we used $\chi_{rA} = r \cdot \chi_A$, since $r \cdot \chi_A(s) = 1$ if and only if $r^{-1}s \in A$, if and only if $s \in rA$.) □

We can use Exercise 5.0.5 to define amenable groups in general. For a locally compact group G we denote by $\mathfrak{B}(G)$ the *Borel σ-algebra on G*, that is, the smallest σ-algebra containing all open subsets of G.

Definition 5.0.6. *Let G be a locally compact group. We say that G is* amenable *if there is a positive linear functional $I : L^\infty(G, \mathfrak{B}(G)) \to \mathbb{C}$ which is positive, has norm 1, and satisfies $I(g \cdot \phi) = I(\phi)$ for all $g \in G$ and all $\phi \in L^\infty(G, \mathfrak{B}(G))$. The functional I is usually called a* left-invariant mean*.*

5.1 Examples of amenable groups

Proposition 5.1.1. *A compact group is amenable.*

Proof. Let μ denote the left Haar measure on a compact group G which is normalised so that $\mu(G) = 1$. Since μ is left-invariant, the functional I_μ on $L^\infty(G, \mathfrak{B}(G))$ is a left-invariant mean. □

In particular, finite groups are amenable.

It is surprisingly difficult to prove groups are amenable. We state the following classic result, due to Markov and Kakutani, to help us sketch the proof of the following result.

Theorem 5.1.2. *Let E be a (locally) convex topological vector space and let $C \subset E$ be compact and convex. Suppose that $(T_i)_{i\in I}$ is a family of maps $T_i : C \to C$ which are linear (even affine suffices) and mutually commuting. Then there is a point $c \in C$ which is a fixed point of every T_i.*

Now we can prove that another class of groups are amenable.

Theorem 5.1.3. *Abelian locally compact groups are amenable.*

Proof. Let $M \subset L^\infty(G)^*$ be the collection of all means on $L^\infty(G)$, that is, the collection of all contractive, positive linear maps $L^\infty(G) \to \mathbb{C}$. The set M is convex and compact. For each $g \in G$ define

$$T_g : L^\infty(G)^* \to L^\infty(G)^*;\ (T_g m)(\phi) := m(g \cdot \phi), \quad m \in L^\infty(G)^*,\ \phi \in L^\infty(G).$$

The maps T_g are linear, (weak*-)continuous, satisfy $T_g(M) \subset M$ and $T_g T_h = T_{gh} = T_{hg} = T_h T_g$ $(g, h \in G)$. By Theorem 5.1.2 there is $m \in M$ with $T_g(m) = m$ for all $g \in G$. This m is a left-invariant mean. □

You might guess that if G is amenable and H is a closed subgroup of G then H is also amenable — just restrict the left-invariant mean on $L^\infty(G)$ to $L^\infty(H)$, right? It turns out that this does not work, because Haar measure on H may not be the restriction of a Haar measure on G. This difficulty can be overcome, but we do not give the proof. The following result collects some hereditary properties of amenability.

Theorem 5.1.4. *Amenability is preserved by the following constructions:*

(i) passing to closed subgroups;

(ii) passing to quotients;

(iii) passing to extensions (in particular, finite direct products of amenable groups are amenable);

(iv) taking increasing unions.

Exercise 5.1.5. *A group G is called* solvable *if there exist subgroups $\{e\} = G_0, G_1, \dots, G_{n-1}, G_n = G$ such that G_i is a normal subgroup of G_{i+1} and G_{i+1}/G_i is abelian for each i. Show that solvable groups are amenable.*

Hint. Apply Theorem 5.1.4 part (iii) repeatedly.

Solution. By Theorem 5.1.3, G_1 is amenable. Now if G_i is amenable then, by definition of solvability, G_{i+1} is an extension of G_i by G_{i+1}/G_i, and G_{i+1}/G_i is abelian and therefore also amenable, so by Theorem 5.1.4 part (iii) G_{i+1} is amenable. Continuing in this way we find $G_n = G$ is amenable. □

Example 5.1.6. *The Euclidean motion groups $\mathbb{E}^1$ and $\mathbb{E}^2$ are solvable (when equipped with the discrete topology), hence amenable.*

Exercise 5.1.7. *Show that the Euclidean motion groups $\mathbb{E}^n$ are not amenable as discrete groups when $n \geq 3$.*

Hint. Apply Theorem 5.1.4 part (i).

Solution. It follows from Theorem 3.1.1 that $\mathbb{E}^3$ (and therefore $\mathbb{E}^n$ with $n \geq 3$) contains a subgroup isomorphic to $\mathbb{F}_2$. When we equip $\mathbb{E}^3$ with the discrete topology this subgroup becomes a closed non-amenable subgroup of $\mathbb{E}^3$, so $\mathbb{E}^3$ cannot be amenable by Theorem 5.1.4 part (i). □

In the above results we regarded $\mathbb{E}^n$ as having the discrete topology, though they also carry a different topology as subgroups of $\mathrm{GL}(n, \mathbb{R})$. On this topic we quote the following result; see [5, Corollary 1.1.10] for a proof.

Proposition 5.1.8. *Let G be a locally compact group. If G is amenable when equipped with the discrete topology then G is amenable with its original topology.*

Consider the Euclidean motion group $\mathbb{E}^n$, which contains two important subgroups: the subgroup of translations, which we identify with $\mathbb{R}^n$, and the subgroup of rotations about some axis through the origin, given by $\mathrm{SO}(n, \mathbb{R})$. In fact, every element of $\mathbb{E}^n$ can be written as a translation by $a \in \mathbb{R}^n$ followed by a rotation by $T \in \mathrm{SO}(n, \mathbb{R})$, say

$$x \mapsto T(x + a), \quad x \in \mathbb{R}^n,$$

or equivalently as a rotation followed by a translation

$$x \mapsto Tx + b, \quad x \in \mathbb{R}^n,$$

with $b = Ta$. The collection of translations forms a normal subgroup of $\mathbb{E}^n$: for any rotation $T \in \mathrm{SO}(3, \mathbb{R})$ the element of $\mathbb{E}^n$ given by $x \mapsto T(T^{-1}(x) + a)$ is obviously again a translation, this time by Ta. This means that $\mathbb{E}^n$ is the semidirect product formed by $\mathrm{SO}(n, \mathbb{R})$ acting on $\mathbb{R}^n$, $\mathbb{E}^n = \mathbb{R}^n \rtimes \mathrm{SO}(n, \mathbb{R})$, or what we called the *extension* of $\mathbb{R}^n$ by $\mathrm{SO}(n, \mathbb{R})$ in Theorem 5.1.4. This theorem then allows us to deduce amenability of $\mathbb{E}^n$ in the Euclidean topology. Indeed, $\mathbb{R}^n$ is abelian and therefore amenable, while $\mathrm{O}(n, \mathbb{R}) = \pi^{-1}(\{I_n\})$, where I_n is the $n \times n$ identity matrix, $\mathrm{O}(n, \mathbb{R}) = \{T \in \mathrm{GL}(n, \mathbb{R}) : T \text{ is orthogonal}\}$, and

$$\pi : \mathrm{GL}(n, \mathbb{R}) \to \mathrm{GL}(n, \mathbb{R}); \; \pi(S) := S^t S.$$

Since π is continuous this means that $\mathrm{O}(n, \mathbb{R})$ is closed, and $\mathrm{O}(n, \mathbb{R})$ is obviously bounded, so by the Heine–Borel Theorem $\mathrm{O}(n, \mathbb{R})$ is compact; now

$\mathrm{SO}(n,\mathbb{R})$ is closed in $\mathrm{O}(n,\mathbb{R})$, hence is compact and therefore amenable. It follows from Theorem 5.1.4 part (iii) that $\mathbb{E}^n$ is amenable.

The message from this section is that, though Euclidean motion groups are amenable in their Euclidean topology, it is non-amenability of $\mathbb{E}^n$ as a discrete group (when $n \geq 3$) that gives rise to the Banach–Tarski paradox.

5.2 Amenability and paradoxical decompositions

It turns out that the notion of amenability is what we needed to solve the problem of measure, which is our final goal. The proof below requires the difficult Hahn–Banach Theorem from functional analysis; it is based on [10, page 161, Theorem 10.11 (i) $\Longrightarrow$ (v)].

Theorem 5.2.1. *Let G be an amenable group of rigid motions of $\mathbb{R}^n$. Then there is a finitely additive, G-invariant extension of Lebesgue measure to all subsets of $\mathbb{R}^n$.*

Proof. Define spaces

$$V_0 := \{\phi : \mathbb{R}^n \to \mathbb{R} :\ \phi \text{ is Lebesgue integrable}\}$$

and

$$V := \{\phi : \mathbb{R}^n \to \mathbb{R} :\ \text{there is } \psi \in V_0 \text{ with } \phi(x) \leq \psi(x) \text{ for all } x \in \mathbb{R}^n\}.$$

One can see that V_0 and V are ($\mathbb{R}$-)linear spaces and V_0 is a subspace of V. Both V and V_0 have actions of G: for ϕ in V or V_0

$$(r \cdot \phi)(x) := \phi(r^{-1} \cdot x) \quad r \in G,\ x \in \mathbb{R}^n.$$

It is clear that $r \cdot \phi \in V_0$ when $\phi \in V_0$: since G acts by isometries it preserves the open, therefore the Lebesgue measurable, sets. If $\phi \in V$ is bounded by $\psi \in V_0$ then clearly $r \cdot \phi$ is bounded by $r \cdot \psi$, so $r \cdot \phi \in V$.

Let $F_0 : V_0 \to \mathbb{R}$ denote the linear map given by Lebesgue integration:

$$F_0(\phi) := \int_{\mathbb{R}^n} \phi(x)\, d\lambda(x), \quad \phi \in V_0.$$

Define a G-invariant sublinear functional $p : V \to \mathbb{R}$ by

$$p(\phi) := \inf\{F_0(\psi) :\ \psi \in V_0 \text{ and } \phi(x) \leq \psi(x) \text{ for all } x \in \mathbb{R}^n\}.$$

Clearly $p(c\phi) = cp(\phi)$ for $c \in \mathbb{R}$, since F_0 and inf both have this property; subadditivity of p follows from the properties of inf and linearity of F_0. Moreover, p is G-invariant:

$$\begin{aligned} p(r \cdot \phi) &= \inf\{F_0(\psi) : \ \psi \in V_0 \text{ and } (r \cdot \phi)(x) \le \psi(x) \text{ for all } x \in \mathbb{R}^n\} \\ &= \inf\{F_0(r^{-1} \cdot \psi) : \ r^{-1} \cdot \psi \in V_0 \text{ and } \phi(x) \le (r^{-1} \cdot \psi)(x) \ \forall\ x \in \mathbb{R}^n\} \\ &= \inf\{F_0(\psi) : \ r^{-1} \cdot \psi \in V_0 \text{ and } \phi(x) \le (r^{-1} \cdot \psi)(x) \ \forall\ x \in \mathbb{R}^n\} \\ &= p(\phi). \end{aligned}$$

By definition $F_0(\phi) \le p(\phi)$ for all $\phi \in V_0$, so by the Hahn–Banach Theorem, Theorem C.0.2, there is a linear map $F : V \to \mathbb{R}$ which extends F_0: $F(\phi) = F_0(\phi)$ for $\phi \in V_0$ and F is dominated by p:

$$-p(-\phi) \le F(\phi) \le p(\phi), \quad \phi \in V.$$

We want to define our extension of Lebesgue measure using this map F, but F is not G-invariant. This is where we must use amenability of G.

Given $\phi \in V$ define another function

$$\theta_\phi : G \to \mathbb{R} \cup \{\infty\}; \ \theta_\phi(r) := F(r^{-1} \cdot \phi), \quad r \in G.$$

We have

$$\begin{gathered} \theta_\phi(r) = F(r^{-1} \cdot \phi) \le p(r^{-1} \cdot \phi) = p(\phi), \\ \theta_\phi(r) = F(r^{-1} \cdot \phi) \ge -p\big(-(r^{-1} \cdot \phi)\big) = -p\big(r^{-1} \cdot (-\phi)\big) = -p(-\phi). \end{gathered}$$

Note that we used linearity of the action of G in the second calculation. Let ν be a measure on G arising from amenability of G, so ν is finitely additive, G-invariant, defined on all subsets of G and $\nu(G) = 1$. Finally, define a measure on $\mathbb{R}^n$ by

$$\mu(A) := \begin{cases} \int_G \theta_{\chi_A}(r)\, d\nu(r) & \text{if } \chi_A \in V; \\ \infty & \text{othwewise.} \end{cases}$$

It remains to prove that this is the measure we are looking for: it is finitely additive, G-invariant, defined on all subsets of $\mathbb{R}^n$ and extends Lebesgue measure. Clearly μ is defined on all subsets of $\mathbb{R}^n$. For finite additivity suppose $A, B \subset \mathbb{R}^n$ are disjoint, and assume $\chi_{A\cup B} \in V$; then

$$\begin{aligned} \mu(A \cup B) &= \int_G \theta_{\chi_{A\cup B}}(r)\, d\nu(r) = \int_G F\big(r^{-1} \cdot (\chi_{A\cup B})\big)\, d\nu(r) \\ &= \int_G F\big(r^{-1} \cdot \chi_A\big) + F\big(r^{-1} \cdot \chi_B\big)\, d\nu(r) = \mu(A) + \mu(B), \end{aligned}$$

where we use that $\chi_{A\cup B} = \chi_A + \chi_B$ (since $A \cap B = \emptyset$), linearity of the action of G and linearity of F. If $\mu(A) = \infty$ or $\mu(B) = \infty$ then clearly $\chi_{A\cup B} \notin V$, so $\mu(A \cup B) = \infty$. For G-invariance, again suppose $\chi_A \in V$, and $r \in G$, so

$$\begin{aligned}\mu(s \cdot A) &= \int_G F(r^{-1} \cdot \chi_{s\cdot A})\, d\nu(r) = \int_G F(\chi_{(r^{-1}s)\cdot A})\, d\nu(r) \\ &= \int_G F\big((s^{-1}r)^{-1} \cdot \chi_A\big)\, d\nu(r) = \int_G F\big(t^{-1} \cdot \chi_A\big)\, d\nu(t) = \mu(A).\end{aligned}$$

We used that $r^{-1} \cdot \chi_B = \chi_{r^{-1}B}$ and also the G-invariance of ν to change the variable. To see that μ extends Lebesgue measure suppose $A \subset \mathbb{R}^n$ and $\lambda(A) < \infty$, then $\chi_A \in V_0$, so

$$\begin{aligned}\mu(A) &= \int_G \theta_{\chi_A}(r)\, d\nu(r) = \int_G F(r^{-1} \cdot \chi_A)\, d\nu(r) = \int_G F_0(r^{-1} \cdot \chi_A)\, d\nu(r) \\ &= \int_G F_0(\chi(A))\, d\nu(r) = \lambda(A) \int_G d\nu(r) = \lambda(A)\nu(G) = \lambda(A).\end{aligned}$$

On the other hand, if $\lambda(A) = \infty$ then $\chi_A \notin V_0$, so $\mu(A) = \infty$ also. We have therefore given the desired extension of Lebesgue measure. □

The above result solves the problem of measure for $\mathbb{R}^n$.

Corollary 5.2.2. *The problem of measure has a positive solution for $\mathbb{R}^1$ and $\mathbb{R}^2$ and a negative solution when $n \geq 3$. More specifically, when $n = 1, 2$ there is an extension of Lebesgue measure to a $\mathbb{E}^n$-invariant measure on all subsets of $\mathbb{R}^n$; when $n \geq 3$ no such extension exists, but one can obtain an extension which is invariant under the action of any amenable subgroup of $\mathbb{E}^n$.*

Proof. We have seen in Example 5.1.6 and Exercise 5.1.7 that $\mathbb{E}^1$ and $\mathbb{E}^2$ are amenable (when regarded as discrete groups) and that $\mathbb{E}^n$ is not amenable for $n \geq 3$ (when regarded as a discrete group it contains a closed non-amenable subgroup isomorphic to $\mathbb{F}_2$). The statements are then immediate from Theorem 5.2.1. □

Appendix A

Integration against finitely-additive measures

Here we give the required definition for integrating against a finitely additive measure.

The *characteristic function* of a set $A \subset X$ is defined

$$\chi_A : X \to \mathbb{C};\ \chi_A(x) := \begin{cases} 1 & \text{if } x \in A; \\ 0 & \text{otherwise.} \end{cases}$$

Recall that an *algebra on* X is a collection of subsets $\mathcal{A}$ of X which contains the empty set and is closed under complements and finite unions (therefore also finite intersections).

Definition A.0.1. *Let $(X, \mathcal{A})$ be a set equipped with an algebra $\mathcal{A}$. A* simple function *from X to $\mathbb{C}$ is a function of the form $\sum_{i=1}^{n} c_i \chi_{A_i}$, where $c_i \in \mathbb{C}$ and $A_i \in \mathcal{A}$.*

There is a natural definition of integral for simple functions.

Definition A.0.2. *Let μ be a finitely-additive measure on $(X, \mathcal{A})$. For a simple function $f = \sum_{i=1}^{n} c_i \chi_{A_i}$ we define the* integral of f with respect to μ *by*

$$\int_X f(x)\, d\mu(x) := \sum_{i=1}^{n} c_i \mu(A_i).$$

This definition, together with the next result, allows us to define the integral of a function. Of course, the integral may take the value $+\infty$. Let

$L^\infty(X, \mu)$ denote the collection of measurable functions $\phi : X \to \mathbb{C}$ such that $\|\phi\|_\infty$ is finite, where $\|\cdot\|_\infty$ denotes the *supremum norm*

$$\|\phi\|_\infty := \operatorname{esssup}\{|\phi(x)| : x \in X\}.$$

We met $L^\infty(X, \mu)$ in Chapter 5, in the case that μ is counting measure, and wrote $\ell^\infty(X)$ in this case.

Proposition A.0.3. *The collection of simple functions is dense in $L^\infty(X, \mu)$ when the latter space is equipped with the supremum norm; that is, for any $\phi \in L^\infty(X, \mu)$ and $\epsilon > 0$ there is a simple function f with $\|f - \phi\|_\infty < \epsilon$.*

Finally, we can define the integral of a function.

Theorem A.0.4. *For a function $\phi \in L^\infty(X, \mu)$ and a finitely-additive measure μ on $(X, \mathcal{A})$ define*

$$\int_X \phi(x)\, d\mu(x) := \lim_k \int_X f_k(x)\, d\mu(x), \quad \|f_k - \phi\|_\infty \xrightarrow{k} 0.$$

The map

$$I_\mu : L^\infty(X, \mathcal{A}) \to \mathbb{C};\ I_\mu(\phi) := \int_X \phi(x)\, d\mu(x)$$

is linear, and positive when μ is.

Appendix B

Haar measure

Definition B.0.1. *A* topological group *is a group G which is also a topological space, such that the operations*

$$G \times G \to G;\ (g,h) \mapsto gh \quad \textit{and} \quad G \to G;\ g \mapsto g^{-1}$$

are continuous. A locally compact group *is a topological group which is locally compact and Hausdorff as a topological space.*

Examples B.0.2. *(i) Any group equipped with the discrete topology (e.g. $\mathbb{Z}$ or $\mathbb{F}_n$) is an example; such groups are called* discrete groups*.*

(ii) The groups $\mathbb{R}^n$ with the Euclidean topology are locally compact groups.

(iii) The matrix groups $\mathrm{GL}(n,\mathbb{R})$ are locally compact with the topology they inherit as a subset of $\mathbb{R}^{n^2}$.

(iv) The group $\mathbb{Q}$ is not a locally compact group with the topology as a subset of $\mathbb{R}$.

Often when working with topological groups it suffices to consider neighbourhoods of the unit element $e \in G$, since if U is an open neighbourhood containing e then gU is an open neighbourhood containing $g \in G$.

We like to work with locally compact groups because they always carry a measure, called *Haar measure*, which interacts well with the group structure. We refer to [2, Section 9.3] for the proof of the following result. Recall that $\mathfrak{B}(G)$ denotes the collection of Borel sets of the topological space G; that is, the smallest σ-algebra on G containing all open subsets of G.

Theorem B.0.3. *Let G be a locally compact group. There is a non-zero regular (countably additive) measure μ on $(G, \mathfrak{B}(G))$ which is* left-invariant*:*

$$\mu(gA) = \mu(A) \quad \text{for all } g \in G,\ A \in \mathfrak{B}(G).$$

Such a measure μ is called a Haar measure*. This measure is unique up to a positive constant, that is, if ν is another Haar measure on $(G, \mathfrak{B}(G))$ then there is $c > 0$ such that $\mu = c\nu$.*

For any discrete group counting measure is a Haar measure, since $|gA| = |\{ga : a \in A\}| = |A|$. When a group is compact it is often convenient to *normalise* Haar measure by choosing the number c in Theorem B.0.3 so that the measure of the whole group is 1. Lebesgue measure on $\mathbb{R}^n$ is also an example of Haar measure (this is really what Theorem 4.1.6 says), but this time we choose c so that $[0, 1]^n$ has measure 1.

Appendix C

The Hahn–Banach Theorem

The Hahn–Banach Theorem is an essential result in functional analysis. We only need a few definitions and the statement of the result; the proof is far beyond the scope of these notes — it can be found in most textbooks on functional analysis.

Definition C.0.1. *Let V be a real vector space. Recall that a* linear functional *on V is a linear map from V to $\mathbb{R}$. A* sublinear functional *on V is a map $p : V \to \mathbb{R}$ such that:*

(i) $p(cv) = cp(v)$ for all $v \in V$ and $c \in [0, \infty)$;

(ii) $p(v + w) \leq p(v) + p(w)$ for all $v, w \in V$.

For example, if $V = \mathbb{R}^n$ the usual Euclidean distance $p(x) := |x|$ is a sublinear functional on $\mathbb{R}^n$.

Now we can state the Hahn–Banach Theorem.

Theorem C.0.2. *Let V be a real vector space and $V_0 \subset V$ a subspace. Suppose that $F_0 : V_0 \to \mathbb{R}$ is a linear functional and $p : V \to \mathbb{R}$ is a sublinear functional such that $F_0(v) \leq p(v)$ for all $v \in V_0$. Then there is a linear functional $F : V \to \mathbb{R}$ which extends F_0:*

$$F(v) = F_0(v) \quad \textit{for all } v \in V_0,$$

and satisfying

$$-p(-v) \leq F(v) \leq p(v) \quad \textit{for all } v \in V.$$

The remarkable thing about the Hahn–Banach Theorem is that we can extend F_0 to a (possibly much larger) space *while still keeping the extension dominated by p.*

Note that the axiom of choice is required to prove the Hahn–Banach Theorem, so this is another place in these notes where the axiom of choice is used in an essential way. The use of the axiom of choice means that the Hahn–Banach Theorem is non-constructive — the only information we have is that the extension exists.

Bibliography

[1] Stafan Banach and Alfred Tarski. Sur la décomposition des ensembles de points en parties respectivement congruentes. *Fund. Math.*, 6:244–277, 1924.

[2] Donald L. Cohn. *Measure theory.* Birkhäuser Advanced Texts: Basler Lehrbücher. [Birkhäuser Advanced Texts: Basel Textbooks]. Birkhäuser/Springer, New York, second edition, 2013.

[3] Felix Hausdorff. Bemerkung über den Inhalt von Punktmengen. *Math. Ann.*, 75(3):428–433, 1914.

[4] Henri Lebesgue. *Intégrale, Longueur, Aire.* PhD thesis, University of Nancy, 1902.

[5] Volker Runde. *Lectures on amenability*, volume 1774 of *Lecture Notes in Mathematics.* Springer-Verlag, Berlin, 2002.

[6] Geoff Smith and Olga Tabachnikova. *Topics in group theory.* Springer Undergraduate Mathematics Series. Springer-Verlag London, Ltd., London, 2000.

[7] Alfred Tarski. Algebraische Fassung des Maßproblems. *Fundam. Math.*, 31:47–66, 1938.

[8] Giuseppe Vitali. Sul problema della misura dei gruppi di punti di una retta. *Tip. Gamb. Parm.*, 1905.

[9] John von Neumann. Zur allgemeinen theorie des maßes. *Fund. Math.*, 13:73–116, 1929.

[10] Stan Wagon. *The Banach–Tarski Paradox.* Encyclopedia of Mathematics and its Applications. Cambridge University Press, 1985.

[11] Tom Weston. The Banach–Tarski paradox. Available online at https://people.math.umass.edu/~weston/oldpapers/banach.pdf.