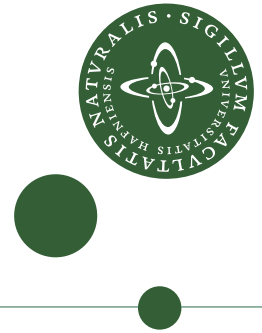


DEPARTMENT OF MATHEMATICAL SCIENCES
UNIVERSITY OF COPENHAGEN

June 12, 2009



The Banach-Tarski Paradox

Søren Knudby

Bachelor Thesis in Mathematics

Advisor: Mikael Rørdam

© Søren Knudby, 2009

Layout by the author using L^AT_EX.
Thanks to Jérôme Baltzersen and
Niels Woo-Sang Kjærsgaard for help
with layout.

Abstract	ii
Introduction	iii
1 The Banach-Tarski paradox	1
1.1 Paradoxicality	1
1.2 Free groups	3
1.3 The Hausdorff paradox	4
1.4 Equidecomposability	6
1.5 The Banach-Tarski paradox	8
2 Tarski's theorem	12
2.1 A cancellation law	13
2.2 Tarski's theorem	15
3 The absence of paradoxes	20
3.1 Measures in groups	20
3.2 Some amenable groups	21
3.3 The real line and the plane	25
A The isometry groups of the Euclidean spaces	29
B The Hahn-Banach Theorem	31
C Integration with finitely additive probability measures	32

This thesis describes the Banach-Tarski paradox and related subjects. The thesis contains two main results: The Banach-Tarski paradox and Tarski's theorem, described in chapter 1 and chapter 2, respectively. The concept of paradoxicality is defined and used throughout the thesis. A number of propositions concerning this concept are proved and used to show the Banach-Tarski paradox – both in its original form and in a stronger version. Tarski's theorem links paradoxicality to measure theory and gives a measure theoretical criterion for, when a paradoxical decomposition exists. In chapter 3, an application of the Hahn-Banach theorem shows that the Banach-Tarski paradox can not be transferred to the plane. In connection with this we will study the class of amenable groups, and a number of fundamental results on these will be proved.

Resumé på dansk (Abstract in Danish)

Dette bachelorprojekt beskriver Banach-Tarskis paradoks og beslægtede emner. Projektet indeholder to hovedresultater: Banach-Tarskis paradoks og Tarskis sætning, som beskrives i henholdsvis kapitel 1 og kapitel 2. Begrebet paradoksalitet defineres og benyttes gennem resten af projektet. En række sætninger vedrørende dette begreb bevises og bruges til at vise Banach-Tarskis paradoks – både i dets oprindelige form og i en stærkere version. Tarskis sætning knytter paradoksalitet til målteori og giver et målteoretisk kriterium for, i hvilke tilfælde paradoksale opspaltninger findes. I kapitel 3 anvendes Hahn-Banachs sætning til at bevise, at Banach-Tarskis paradoks ikke kan overføres til planen. I denne forbindelse indføres de amenable grupper, og en række grundlæggende resultater om disse bevises.

Through history the concept of infinity has puzzled mankind and maybe particularly mathematicians. An early example of this is Galilei's attempt to compare the size of the set of square numbers with the size of the set of all natural numbers. By using different approaches he was justified in answering both yes and no to the question: Are there more natural numbers than square numbers? This was of course most unsatisfactory. Paradoxes like that show us how careful we must be, if we want to work with the concept of infinity in a satisfactory manner, and how the concept often leads to results that may seem paradoxical.

With the development of Cantor's set theory and the concept of cardinality, mathematicians found a way to deal with the infinite. But results that seem counter-intuitive still find their way into the mathematical world. Often such results are related to the *Axiom of Choice*, which was formulated in the early 20th century by Zermelo. It has several equivalent formulations, one of which is the following:

Axiom of Choice.

If $\{A_i\}_{i \in I}$ is a collection of disjoint sets, it is possible to find a set C , such that C contains exactly one element from each of the sets A_i , $i \in I$.

This seemingly harmless axiom has many applications in modern mathematics, but has also brought with it a number of highly counter-intuitive results. The Banach-Tarski paradox is one example of this, but before stating it we should be clear about one thing: The Banach-Tarski paradox is *not* a paradox in the usual sense of the word. It is simply a theorem which at first seems false, but nevertheless can be proved rigorously. It may be formulated as

The Banach-Tarski paradox.

There exists a partition of the unit ball B into finitely many pieces A_1, \dots, A_n , B_1, \dots, B_m and isometric mappings $\varphi_1, \dots, \varphi_n, \psi_1, \dots, \psi_m$ such that

$$B = \bigcup_{i=1}^n \varphi_i(A_i) = \bigcup_{j=1}^m \psi_j(B_j).$$

Loosely, the Banach-Tarski paradox may be formulated as: an orange may be cut into a finite number of pieces and rearranged to collect two oranges of the same size as the original. Of

course the theorem questions whether the axiom of choice is a reasonable axiom to include or not. There has been a long discussion about the use of the axiom of choice in mathematics, but most mathematicians agree that it should be part of the standard axioms. We shall not go into this discussion.

The Banach-Tarski paradox is named after the two mathematicians Stefan Banach and Alfred Tarski, who proved the theorem in 1924. Their work was related to measure theory, which was developed in the early 20th century. In 1905 Vitali showed the existence of a non-measurable set and thereby proved the non-existence of a (countably additive) extension of the Lebesgue measure to all subsets of the real line. In 1914 Hausdorff proved the non-existence of a certain kind of measure by constructing a paradox related to the sphere. These are some of the results that motivated Banach and Tarski.

The strong connection between paradoxicality and measure theory will not be emphasized until chapter two. The reader is supposed to be familiar with elementary group theory and in particular group actions. Chapter two and especially chapter three use basic measure theory, and a few proofs in these chapters also use the Hahn-Banach theorem, in connection with which some acquaintance with functional analysis may be helpful. Chapter three also uses general topology and Tychonoff's theorem in a single proof.

I wish to thank my advisor Mikael Rørdam for his ideas and advice.

1.1 Paradoxicality

Usually the Banach-Tarski paradox is formulated through the concept of paradoxicality, which will be defined below. Paradoxicality is a way of formalizing the idea of duplicating a set by splitting it into pieces and moving the pieces around. Group actions are a natural way of moving the pieces, so first recall the definition of a group action. A group G is said to act on a set X , if there is a mapping from $G \times X$ to X , denoted $(g, x) \mapsto g.x$, such that for all $x \in X, g, h \in G$

$$1.x = x, \quad g.(h.x) = (gh).x$$

where 1 denotes the neutral element of G . To each $g \in G$ is associated a map $x \mapsto g.x$, and we will denote this map also by g . This should not cause any confusion. Often we will omit the dot and just write $g(x)$ or simply gx instead of $g.x$.

Example 1.1. The group \mathbb{E}_n of isometries of the Euclidean space \mathbb{R}^n acts on \mathbb{R}^n . The subgroup SO_n of orientation-preserving orthogonal transformations also acts on \mathbb{R}^n . SO_n is called the rotation group. The unit sphere S^{n-1} is stable under this action, so SO_n also acts on S^{n-1} .

Example 1.2. Any group acts on itself by left multiplication, and a subgroup also acts on the whole group by left multiplication. A group also acts on its power set $\mathcal{P}(G)$ by left multiplication, and if Y is any set, then G acts on Y^G (the set of maps from G to Y), by letting $g.f$ be the map $h \mapsto f(g^{-1}h)$. Here $g, h \in G$ and $f \in Y^G$.

The reader is supposed to be familiar with these standard actions. We will now formulate the concept of paradoxicality.

Definition 1.3. Let G be a group acting on a set X , and let $E \subseteq X$. If there exist disjoint subsets $A_1, \dots, A_n, B_1, \dots, B_m$ of E and group elements $g_1, \dots, g_n, h_1, \dots, h_m$ such that

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

1.1 Paradoxicality

then we say that E is G -paradoxical. If it is obvious which group is meant we simply say that X is paradoxical, or if $X = G$ and the action is left multiplication we will also just say that G is paradoxical.

Remark 1.4. Let us make a few comments on this definition since it is the most important definition in the thesis. It is not required that the pieces $\{A_i\}_{i=1}^n \cup \{B_j\}_{j=1}^m$ cover all of E . It will nevertheless be shown later (corollary 1.23) that it is always possible to choose the pieces so that they cover E .

There are no restrictions on how the pieces may look, eg. measurability, connectedness, etc. That is one of the reasons why some results in connection with paradoxicality may seem counter-intuitive.

Also note that we only allow finitely many pieces.

With the terminology of paradoxicality in place we can formulate the Banach-Tarski paradox as follows:

Theorem 1.5 (The Banach-Tarski paradox). *The unit ball in \mathbb{R}^3 is \mathbb{E}_3 -paradoxical.*

The proof will be postponed until section 1.5. We will now show how the paradoxicality of a group G can be used to produce a paradoxical decomposition of a set X , if G acts on X . This is not always the case, but if the action is *free*, i.e. only the neutral element of G has fixed points, it can be done. Recall that if $g \in G \setminus \{1\}$ we say that x is a non-trivial fixed point if $g.x = x$.

Proposition 1.6. *If G acts on X without non-trivial fixed points and G is paradoxical, then X is G -paradoxical.*

Proof. Let $A_1, \dots, A_n, B_1, \dots, B_m \subseteq G, g_1, \dots, g_n, h_1, \dots, h_m \in G$ witness the paradoxicality of G . If X/G denotes the orbit space it is possible to choose exactly one element from each orbit (this requires the axiom of choice). If we collect these representatives in a set M then $\{g.M \mid g \in G\}$ is a partition of X . To see this assume first that $g.M \cap h.M \neq \emptyset$. Then $g.x = h.y$ for some $x, y \in M$. Hence $h^{-1}g.x = y$, so x and y belong to the same orbit, and we see $x = y$ by the choice of M . Since the action is free, and x is fixed by $h^{-1}g$, we get $h^{-1}g = 1$, or $h = g$, so $g.M = h.M$.

To see that $\{g.M \mid g \in G\}$ covers all of X suppose that $x \in X$ is given. M contains an element from each orbit, so let $y \in G.x \cap M$ be such an element. Then $g.y = x$ for some $g \in G$ and hence $x \in g.M$.

Now, for $1 \leq i \leq n, 1 \leq j \leq m$ define

$$A_i^* = \bigcup_{g \in A_i} g.M = A_i.M, \quad B_j^* = \bigcup_{g \in B_j} g.M = B_j.M.$$

Then clearly $A_1^*, \dots, A_n^*, B_1^*, \dots, B_m^*$ are disjoint subsets of X . Further

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

and in a similar way we get

$$\bigcup_{j=1}^m h_j.B_j^* = X$$

showing that X is G -paradoxical. □

Since a subgroup acts without nontrivial fixed points on the whole group we get the following immediate corollary:

Corollary 1.7. *If a group G has a paradoxical subgroup then G itself is paradoxical.*

The following proposition is a converse of proposition 1.6, but will not be important (until chapter 3). We provide it only for completeness and to work with the concept of paradoxicality.

Proposition 1.8. *If X is G -paradoxical, then G is paradoxical.*

Proof. Let $A_1, \dots, A_n, B_1, \dots, B_m \subseteq X$, $g_1, \dots, g_n, h_1, \dots, h_m \in G$ witness the paradoxicality of X with respect to G . Fix some $x \in X$, and for $i = 1, \dots, n$, $j = 1, \dots, m$ define

$$A_i^* = \{g \in G \mid g.x \in A_i\} \quad \text{and} \quad B_j^* = \{g \in G \mid g.x \in B_j\}.$$

These are all disjoint, because $A_1, \dots, A_n, B_1, \dots, B_m$ are. The orbit $G.x$ is the union of the sets $g_i.A_i \cap G.x$, where $i = 1, \dots, n$, so if $g \in G$ then $g.x$ is in some $g_i.A_i$. This means that $g.x = g_i.a$ for some i and $a \in A_i$. Then $g_i^{-1}g.x = a \in A_i$. So $g_i^{-1}g$ is in A_i^* , and hence $g \in g_i.A_i^*$, which shows that

$$G = \bigcup_{i=1}^n g_i.A_i^*$$

A similar argument shows that G

$$G = \bigcup_{j=1}^m h_j.B_j^*$$

demonstrating the paradoxicality of G . □

1.2 Free groups

Free groups will be an important source of paradoxical groups. Since the reader is not assumed to be familiar with free groups we introduce the concept. If S is a set, the free group generated by S is the group of all reduced finite words with letters from $\{s, s^{-1} \mid s \in S\}$. A word is

1.3 The Hausdorff paradox

called reduced if it contains no pairs of adjacent letters of the form ss^{-1} or $s^{-1}s$. The group composition is concatenation of words followed by reduction, that is removing pairs of the mentioned forms. Unless otherwise mentioned, the generators are always assumed to have infinite order.

It can be proved that if the sets S and T generate the same free group, then they have the same cardinality. The cardinality of a generating set is called the rank of the free group. A free group is completely determined (up to isomorphism) by its rank, if we assume that all generators have infinite order.

Proposition 1.9. *The free group F_2 of rank two is paradoxical.*

Proof. Let $\{\sigma, \tau\}$ be a generating set for F_2 . For $\rho \in \{\sigma, \sigma^{-1}, \tau, \tau^{-1}\}$ let $W(\rho)$ denote the set of all words beginning (on the left) with ρ . Then F_2 can be written as the disjoint union

$$F_2 = \{1\} \cup W(\sigma) \cup W(\sigma^{-1}) \cup W(\tau) \cup W(\tau^{-1})$$

We claim that $F_2 = W(\sigma) \cup \sigma W(\sigma^{-1})$. To see this suppose that $w \in F_2 \setminus W(\sigma)$. Then $\sigma^{-1}w \in W(\sigma^{-1})$, and from this we infer that $w = \sigma(\sigma^{-1}w) \in \sigma W(\sigma^{-1})$. In a similar way it can be shown that $F_2 = W(\tau) \cup \tau W(\tau^{-1})$, showing that F_2 is paradoxical. \square

The combination of corollary 1.7 and the above proposition yields the following corollary:

Corollary 1.10. *If a group G contains a free subgroup of rank 2, then G is paradoxical.*

1.3 The Hausdorff paradox

As a step on the way to prove the Banach-Tarski paradox we will prove the Hausdorff paradox which gives a paradoxical decomposition of almost all of the unit sphere S^2 . The way to do this is to realize F_2 as a subgroup of SO_3 and then use proposition 1.6 to lift the paradoxical behavior of the subgroup to a major part of the unit sphere. A set S of elements in a group are called independent, if no non-trivial, reduced word using letters from S and their inverses is the identity. Hence a pair of independent elements will generate a free subgroup of rank two.

Lemma 1.11. *There exist two independent rotations about axes through the origin in \mathbb{R}^3 . Hence SO_3 contains a free subgroup of rank 2.*

Proof. Let σ be a counter-clockwise rotation about the z -axis through the angle $\theta = \arccos 3/5$, and let τ be a rotation about the x -axis through the same angle. The rotations are represented by the matrices

$$\sigma = \begin{bmatrix} 3/5 & -4/5 & 0 \\ 4/5 & 3/5 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \tau = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3/5 & -4/5 \\ 0 & 4/5 & 3/5 \end{bmatrix}.$$

We must show that no non-trivial, reduced word in $\{\sigma, \sigma^{-1}, \tau, \tau^{-1}\}$ is the identity. Since conjugation by σ does not change whether or not a word is the identity, we need only consider the case with words ending (on the right) with σ . We claim that for any such word w the point $w(1, 0, 0)$ will be of the form $(a, b, c)/5^n$ where n is the length of w , and a, b, c are integers. Moreover b is not divisible by 5. This will be sufficient because then $w(1, 0, 0) \neq (1, 0, 0)$, so w is not the identity.

The claim is shown by induction over n . For $n = 1$, we have $w = \sigma$ and $w(1, 0, 0) = (3, 4, 0)/5$. Assume now that w is a word of length $n \geq 2$. There are the possibilities

$$w = \sigma^{\pm 1}w' \quad \text{or} \quad w = \tau^{\pm 1}w',$$

where w' is a word of length $n - 1$. If $w'(1, 0, 0) = (a', b', c')/5^{n-1}$, then

$$\sigma^{\pm 1}w'(1, 0, 0) = (3a' \pm 4b', 3b' \mp 4a', 5c')/5^n \tag{1.1}$$

and

$$\tau^{\pm 1}w'(1, 0, 0) = (5a', 3b' \pm 4c', 3c' \mp 4b')/5^n. \tag{1.2}$$

By the induction hypothesis a', b', c' are integers, so a, b, c are integers. We must check that 5 is not a divisor of b . We split this into four cases according to the first two letters of w :

$$w = \sigma^{\pm 1}\tau^{\pm 1}w'', \quad w = \tau^{\pm 1}\sigma^{\pm 1}w'', \quad w = \tau^{\pm 1}\tau^{\pm 1}w'', \quad w = \sigma^{\pm 1}\sigma^{\pm 1}w'',$$

where w'' is any (possibly empty) word. We will only go through the first and the last case, since the remaining two cases are similar to those. The last case only refers to the possibilities $w = \sigma\sigma w''$ and $w = \sigma^{-1}\sigma^{-1}w''$, since w is assumed to be reduced. We remark that $w''(1, 0, 0)$ is of the form $(a'', b'', c'')/5^{n-2}$, where a'', b'', c'' are integers, but we can not assume that b'' is not divisible by 5, since this is not a part of the induction hypothesis.

If $w = \sigma^{\pm 1}\tau^{\pm 1}w''$, then by (1.1) and (1.2) we get $b = 3b' \mp 4a'$ where $a' = 5a''$ and (by the induction hypothesis) 5 does not divide b' . This shows that b is not divisible by 5.

If $w = \sigma^{\pm 1}\sigma^{\pm 1}w''$ then by applying (1.1) twice we see that $a' = 3a'' \pm 4b''$ and $b = 3b' \mp 4a'$. From this

$$\begin{aligned} b &= 3b' \mp 4(3a'' \pm 4b'') = 3b' \mp 12a'' - 16b'' - 9b'' + 9b'' \\ &= 3b' + 3(3b'' \mp 4a'') - 25b'' = 6b' - 25b'', \end{aligned}$$

and since 5 does not divide b' we conclude that b is not divisible by 5. This completes the proof. □

Let F be the subgroup of SO_3 generated by σ and τ . F is paradoxical by proposition 1.9, but in order to apply proposition 1.6 to the action of F on the sphere, the action must be

1.4 Equidecomposability

free. Since any rotation through origo apart from the identity has exactly two fixed points (the intersection of the axis of rotation with the sphere) it is not free. However, if we let

$$D = \{x \in S^2 \mid \rho.x = x \text{ for some } \rho \in F \setminus \{1\}\}, \quad (1.3)$$

then $S^2 \setminus D$ is stable under the action of F . To see this suppose $\rho.x \in D$ where $\rho \in F$. Then $\varphi(\rho.x) = \rho.x$ for some $\varphi \in F$. Hence $\rho^{-1}\varphi\rho.x = x$, so $x \in D$. This shows that $x \in S^2 \setminus D$ implies that $\rho.x \in S^2 \setminus D$. Hence $S^2 \setminus D$ is F -stable, so F acts on $S^2 \setminus D$.

$F \setminus \{1\}$ is countable and each $\rho \in F \setminus \{1\}$ has only two fixed points, so D is countable. Also the action of F on $S^2 \setminus D$ is free, and an application of proposition 1.6 now gives an F -paradoxical decomposition of $S^2 \setminus D$. Since F is a subgroup of SO_3 we have now proved

Theorem 1.12 (The Hausdorff paradox). *There exists a countable subset D of the unit sphere S^2 such that $S^2 \setminus D$ is SO_3 -paradoxical.*

The proof is easily generalized to spheres of arbitrary radius.

1.4 Equidecomposability

We will now introduce the important concept of equidecomposability. It turns out that equidecomposability is an equivalence relation on $\mathcal{P}(X)$, and that G -paradoxicality is a class property with respect to this relation. We will be able to show that S^2 and $S^2 \setminus D$, where D is as in (1.3), belong to the same equivalence class, yielding the SO_3 -paradoxicality of S^2 . Then we will be ready to prove the Banach-Tarski paradox.

Definition 1.13. If G acts on X and A, B are subsets of X we say that A and B are G -equidecomposable and write $A \sim_G B$ if there exist a partition $\{A_i\}_{i=1}^n$ of A and a partition $\{B_i\}_{i=1}^n$ of B and $g_1, \dots, g_n \in G$, such that

$$B_i = g_i.A_i \text{ for every } i \in \{1, \dots, n\}.$$

If \coprod denotes a disjoint union then we may formulate G -equidecomposability as

$$A = \coprod_{i=1}^n A_i, \quad B = \coprod_{i=1}^n B_i, \quad B_i = g_i.A_i \text{ for } i \in \{1, \dots, n\}.$$

Loosely speaking two subsets of X are G -equidecomposable if one of them can be cut into finitely many pieces, and the pieces can be used to build up the other. It is clear from the definition that $A \sim_G B$ if and only if there is a bijection $\gamma : A \rightarrow B$ such that

$$\gamma(x) = \begin{cases} g_1.x & \text{if } x \in A_1 \\ \vdots & \\ g_n.x & \text{if } x \in A_n \end{cases} \quad (1.4)$$

In other words there exist finitely many elements $g_1, \dots, g_n \in G$ and a bijection $\gamma : A \rightarrow B$ such that $\gamma(x) \in \{g_1.x, \dots, g_n.x\}$ for every $x \in A$. Such a bijection will be called a *G-transformation* from A to B . We use the notation $A \sim_\gamma B$ to indicate that $\gamma : A \rightarrow B$ is a *G-transformation* from A to B . We will use this alternative view on equidecomposability many times.

Proposition 1.14. *Let G act on X . The relation \sim_G is an equivalence relation on $\mathcal{P}(X)$.*

Proof. Reflexivity and symmetry are obvious so we will focus on transitivity. Suppose $A \sim_G B$ and $B \sim_G C$. Then there are two *G*-transformations φ and ψ such that $A \sim_\varphi B$ and $B \sim_\psi C$. Then $\psi\varphi : A \rightarrow C$ is a bijection. If $\varphi(x) \in \{g_1.x, \dots, g_n.x\}$ for every $x \in A$ and $\psi(x) \in \{h_1.x, \dots, h_m.x\}$ for every $x \in B$ then

$$(\psi\varphi)(x) \in \{h_j g_i . x \mid j = 1, \dots, m, i = 1, \dots, n\} \text{ for every } x \in A$$

showing that $\psi\varphi$ is a *G*-transformation. Hence $A \sim_{\psi\varphi} C$, so $A \sim_G C$. □

What we have actually shown is that if $\varphi : A \rightarrow B$ and $\psi : B \rightarrow C$ are *G*-transformations, then $\psi\varphi : A \rightarrow C$ is a *G*-transformation.

There is an intimate connection between paradoxicality and equidecomposability. That is the content of the following proposition.

Proposition 1.15. *If G acts on X and $E \subseteq X$, then E is *G*-paradoxical if and only if there are disjoint subsets A and B of E , such that $A \sim_G E \sim_G B$.*

Proof. If $A \sim_G E \sim_G B$ where A and B are disjoint subsets of E , then clearly E is paradoxical. The problem with the other direction is that the pieces used in the definition of paradoxicality may overlap when they have been moved, i.e. the A_i 's are disjoint, but the $g_i.A_i$'s are not necessarily disjoint. But if there are overlaps we can just use smaller pieces to begin with. Let us formalize this idea.

Suppose $A_1, \dots, A_n, B_1, \dots, B_m$ and $g_1, \dots, g_n, h_1, \dots, h_m$ witness the *G*-paradoxicality of E . Define inductively

$$A_1^* = A_1, \quad A_i^* = A_i \setminus g_i^{-1} \left(\bigcup_{k=1}^{i-1} g_k . A_k^* \right) \quad \text{for } i = 2, \dots, n$$

Then $A_i^* \subseteq A_i$ for $i = 1, \dots, n$ so the A_i^* 's are disjoint. Further

$$E = \bigsqcup_{i=1}^n g_i . A_i^*,$$

so if we let

$$A^* = \bigsqcup_{i=1}^n A_i^*,$$

1.5 The Banach-Tarski paradox

then $A^* \sim_G E$. In the same way we may define B^* and show $B^* \sim_G E$. \square

Now we can show that paradoxicality is a class property with respect to equidecomposability.

Proposition 1.16. *If $A \sim_G B$ and A is G -paradoxical, then B is G -paradoxical.*

Proof. Let $\varphi : A \rightarrow B$ be a G -transformation, such that $A \sim_\varphi B$, and let $A_1 \sim_G A \sim_G A_2$, where A_1 and A_2 are disjoint. Since φ is a bijection, $\varphi(A_1)$ and $\varphi(A_2)$ are disjoint, and

$$\varphi(A_1) \sim_G A_1 \sim_G A \sim_G \varphi(A) \quad \text{and} \quad \varphi(A_2) \sim_G A_2 \sim_G A \sim_G \varphi(A),$$

so if we put $B_1 = \varphi(A_1)$ and $B_2 = \varphi(A_2)$ then $B_1, B_2 \subseteq \varphi(A)$ are disjoint. Since $\varphi(A) = B$ we get

$$B_1 \sim_G B \sim_G B_2,$$

and so B is G -paradoxical. \square

1.5 The Banach-Tarski paradox

With equidecomposability well in place we are almost ready to prove the Banach-Tarski paradox. The proof of the following lemma is called a proof of absorption, because it shows how a set that is somehow small in the context of paradoxicality can be ignored.

Lemma 1.17. *If D is a countable subset of S^2 , then $S^2 \setminus D \sim_{\text{SO}_3} S^2$.*

Proof. Let l be a line through the origin that does not intersect the countable set D . For $\theta \in [0, 2\pi[$ we let ρ_θ denote the rotation about l through the angle θ . The orientation of l is immaterial; just choose one of the two possible. For $d \in D, n \in \mathbb{N}$ define

$$A_{d,n} = \{\theta \in [0, 2\pi[\mid \rho_\theta^n(d) \in D\} \quad \text{and} \quad A = \bigcup_{d \in D} \bigcup_{n=1}^{\infty} A_{d,n}$$

Each $A_{d,n}$ is countable, and so is A . Now let $\theta \in [0, 2\pi[\setminus A$. Then $\rho_\theta^n(D) \cap D = \emptyset$ for every natural number n . Also $\rho_\theta^n(D) \cap \rho_\theta^m(D) = \emptyset$ for $n \neq m$, so the sets $D, \rho_\theta(D), \rho_\theta^2(D), \dots$ are pairwise disjoint. If we let

$$\overline{D} = \bigcup_{n=0}^{\infty} \rho_\theta^n(D),$$

then we conclude

$$S^2 = (S^2 \setminus \overline{D}) \cup \overline{D} \sim_{\text{SO}_3} (S^2 \setminus \overline{D}) \cup \rho_\theta(\overline{D}) = S^2 \setminus D.$$

\square

The lemma shows that S^2 and $S^2 \setminus D$ belong to the same equidecomposability class, and since paradoxicality is a class property the Hausdorff paradox shows that S^2 is SO_3 -paradoxical.

Theorem 1.18 (The Banach-Tarski paradox). *The unit ball B in \mathbb{R}^3 is \mathbb{E}_3 -paradoxical. So is any other solid ball in \mathbb{R}^3 .*

Proof. Since S^2 is SO_3 -paradoxical, the radial correspondence $P \mapsto \{\alpha P \mid 0 < \alpha \leq 1\}$ of the unit sphere with the unit ball except origo yields a paradoxical decomposition of $B \setminus \{0\}$. Again we use the trick of absorption to show $B \sim_{\mathbb{E}_3} B \setminus \{0\}$ from which the first assertion will follow.

Let l be the line parallel to the x -axis and through the point $(1/2, 0, 0)$, and let $\rho \in \mathbb{E}_3$ be a rotation about l of infinite order. Put $\bar{D} = \{\rho^n(0) \mid n \geq 0\}$. Then

$$B = (B \setminus D) \cup D \sim_{\mathbb{E}_3} (B \setminus D) \cup \rho(D) = B \setminus \{0\}$$

By using another radial correspondence $P \mapsto \{\alpha P \mid 0 < \alpha \leq r\}$ we get in exactly the same way that any solid ball centered at origo of radius $r > 0$ is \mathbb{E}_3 -paradoxical, and since \mathbb{E}_3 contains all translations, any solid ball is \mathbb{E}_3 -paradoxical. \square

Remark 1.19. If we use the correspondence $P \mapsto \{\alpha P \mid 0 < \alpha\}$ between S^2 and $\mathbb{R}^3 \setminus \{0\}$ we get a paradoxical decomposition of $\mathbb{R}^3 \setminus \{0\}$, and also here $\{0\}$ can be absorbed, so actually \mathbb{R}^3 is \mathbb{E}_3 -paradoxical.

We can improve the result of the theorem even more. But first a definition.

Definition 1.20. If $A, B \subseteq X$ we write $A \preceq_G B$ if A is G -equidecomposable with a subset of B .

Notice that \preceq_G is a class property with respect to G -equidecomposability. The notation suggests that \preceq_G is a partial order on the \sim_G -equivalence classes, and this is indeed the case. Since \preceq_G reflexivity and transitivity are obvious the claim is justified by the following theorem by Banach which uses the idea of the famous Schröder-Bernstein theorem.

Lemma 1.21. *Suppose G acts on X and $A, B, A', B' \subseteq X$.*

- (a) *If $\gamma : A \rightarrow B$ is a G -transformation and $A' \subseteq A$, then $\gamma|_{A'}$ is a G -transformation between A' and $\gamma(A_1)$. In particular $A' \sim_G \gamma(A')$.*
- (b) *If A and A' are disjoint, $\gamma : A \cup A' \rightarrow B \cup B'$, $\gamma|_A : A \rightarrow B$ and $\gamma|_{A'} : A' \rightarrow B'$ are G -transformations, then γ is a G -transformation. In particular $A \cup A' \sim_G B \cup B'$.*

Proof. The first assertion is immediate, since $\gamma|_{A'}$ is also of the form (1.4). Since $\gamma|_A : A \rightarrow B$ and $\gamma|_{A'} : A' \rightarrow B'$ are bijections and $A \cap A' = \emptyset$, then $\gamma : A \cup A' \rightarrow B \cup B'$ is a bijection. It is clear that γ can be written in the form (1.4) with $n + m$ elements from G (sometimes fewer) if $\gamma|_A$ and $\gamma|_{A'}$ can be written in that form with respectively n and m elements from G . \square

1.5 The Banach-Tarski paradox

Theorem 1.22 (Banach-Schröder-Bernstein). *If $A \preceq_G B$ and $B \preceq_G A$, then $A \sim_G B$. Hence \preceq_G is a partial order of the \sim_G -equivalence classes.*

Proof. By assumption there exist G -transformations f and g , such that $A \sim_f B'$ and $B \sim_g A'$ where $A' \subseteq A$ and $B' \subseteq B$. Define $C_0 = A \setminus A' = A \setminus g(B)$ and for $n \in \mathbb{N}$ define inductively $C_n = g(f(C_{n-1}))$. Put

$$C = \bigcup_{n=0}^{\infty} C_n.$$

Then $A \setminus C \subseteq A'$ and $g^{-1}(A \setminus C) = B \setminus f(C)$. By lemma 1.21 (a) we get $A \setminus C \sim_G B \setminus f(C)$ and $C \sim_G f(C)$. We conclude the proof by applying part (b) to get

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

□

As a consequence of the Banach-Schröder-Bernstein theorem we can prove the claim from remark 1.4.

Corollary 1.23. *$E \subseteq X$ is G -paradoxical, if and only if there is a partition of E*

$$E = A \amalg B$$

with $A \sim_G E \sim_G B$. In other words the pieces used in the decomposition of E may be chosen so that they partition E .

Proof. Using the characterization of proposition 1.15 we assume $A \sim_G E \sim_G B$ for disjoint subsets $A, B \subseteq E$. Let $A' = E \setminus B$. Then $A \subseteq A' \subseteq E$, so $A \preceq_G A' \preceq_G E \preceq_G A$, and the Banach-Schröder-Bernstein theorem yields $A' \sim_G E$. Then A' and B are the required pieces. This shows the corollary up to a change of notation. □

It is now time to improve the Banach-Tarski paradox by using the Banach-Schröder-Bernstein theorem.

Theorem 1.24 (Banach-Tarski – strong version). *Any two bounded subsets of \mathbb{R}^3 with non-empty interior are \mathbb{E}_3 -equidecomposable. In particular, any such set is \mathbb{E}_3 -paradoxical.*

Proof. Let A and B be bounded subsets of \mathbb{R}^3 with non-empty interior. By the assumptions we can find solid balls K, L with $A \subseteq K$ and $L \subseteq B$. Choose n so large that K may be covered by n copies of L . This is possible, because K is bounded. Thus if we let M be a union of n disjoint copies of L , $K \preceq_{\mathbb{E}_3} M$. By using the Banach-Tarski paradox repeatedly $L \sim_{\mathbb{E}_3} M$, and hence

$$A \preceq_{\mathbb{E}_3} K \preceq_{\mathbb{E}_3} M \preceq_{\mathbb{E}_3} L \preceq_{\mathbb{E}_3} B$$

showing $A \preceq_{\mathbb{E}_3} B$.

By a similar argument $B \preceq_{\mathbb{E}_3} A$, and the first part of the theorem is now a consequence of the antisymmetry of $\preceq_{\mathbb{E}_3}$.

To prove the second part, let A be a bounded subset of \mathbb{R}^3 with non-empty interior, and let K be an open ball contained in A . K contains two disjoint, open balls K_1 and K_2 , and by the first part of the theorem $K_1 \sim_{\mathbb{E}_3} A \sim_{\mathbb{E}_3} K_2$, so A is \mathbb{E}_3 -paradoxical. \square

We will now set out to prove Tarski's theorem which links paradoxicality to measure theory. It states that a set $E \subseteq X$ is not G -paradoxical if and only if there is a finitely additive measure μ defined on all subsets of X , normalizing E and invariant under the group action. One part of the theorem is trivial, because if E is G -paradoxical there can not exist such a measure: if A and B partition E and $A \sim_G E \sim_G B$, then

$$1 = \mu(E) = \mu(A \cup B) = \mu(A) + \mu(B) = \mu(E) + \mu(E) = 2.$$

The other part of the theorem is the interesting part and requires some work. We start by giving a definition. In all of the following G is a group acting on a set X .

Definition 2.1. Let $E \subseteq X$. A function $\mu : \mathcal{P}(X) \rightarrow [0, \infty]$ satisfying $\mu(\emptyset) = 0$ and $\mu(A \cup B) = \mu(A) + \mu(B)$ if $A \cap B = \emptyset$ for all $A, B \subseteq X$ is called a *finitely additive measure* on X . If $\mu(E) = 1$, we say that μ *normalizes* E , and if $\mu(gA) = \mu(A)$ for all $A \subseteq X, g \in G$, we say that μ is *G -invariant*.

Remark 2.2. If $\mu(E) = 1$, it follows from the finite additivity that $\mu(\emptyset) = 0$, because

$$\mu(E) = \mu(E \cup \emptyset) = \mu(E) + \mu(\emptyset).$$

Subtracting $\mu(E)$, which is possible because $\mu(E)$ is finite, gives $\mu(\emptyset) = 0$.

In the light of the finite additivity, the invariance property may be stated as $\mu(A) = \mu(B)$ if $A \sim_G B$.

Recall that a G -transformation is a bijection $\gamma : A \rightarrow B$ between subsets A and B of X with finitely many $g_1, \dots, g_n \in G$ such that $\gamma(x) \in \{g_1.x, \dots, g_n.x\}$ for every $x \in A$. We note that E is G -paradoxical if and only if there are G -transformations γ_1, γ_2 such that $\gamma_1 : A \rightarrow E$ and $\gamma_2 : B \rightarrow E$ for some disjoint sets $A, B \subseteq E$.

We will also consider several copies of X , and for this purpose we introduce the notation $I_n = \{1, \dots, n\}$ and $E_n = E \times I_n$ for a subset E of X . If S_n denotes the symmetric group on

I_n , then the group $G_n = G \times S_n$ acts on X_n by the componentwise action

$$(g, \sigma).(x, m) = (g.x, \sigma.m),$$

where $(g, \sigma) \in G \times S_n$ and $(x, m) \in X_n$. This is the natural generalization of the action of G on X to an action on n copies of X .

The reason for introducing several copies of X is that we may now formulate paradoxicality as: E is G -paradoxical if and only if $E \times \{1\} \sim_{G_2} E \times \{1, 2\}$, i.e. E is G_2 -equidecomposable with two copies of itself. Another reason for introducing this is that it may often be easier to prove that E_n is G_n -paradoxical for some suitably large $n \in \mathbb{N}$ than to prove directly that E is G -paradoxical. Fortunately this will suffice to prove the G -paradoxicality of E as stated in proposition 2.3.

Note that if E_n is G_n -paradoxical, then E_n is automatically G_m -paradoxical for any natural number $m \geq n$ when we consider E_n as a subset of X_m . Also if E_n is G_m -paradoxical, then E_{kn} is G_m -paradoxical if $kn \leq m$, $k \in \mathbb{N}$.

2.1 A cancellation law

Proposition 2.3. *E is G -paradoxical if and only if E_n is G_n -paradoxical for some $n \in \mathbb{N}$.*

Proof. One direction is trivial. The proof of the other direction is based on the following two lemmas.

Lemma 2.4. *If E is partitioned in two ways, $E = A \amalg B = C \amalg D$, and $A \sim_G B$ and $C \sim_G D$, then $A \sim_G C$.*

Proof. Let f and g be G -transformations witnessing $A \sim_f B$ and $C \sim_g D$. Define $\varphi : E \rightarrow E$ by $\varphi|_A = f$ and $\varphi|_B = f^{-1}$, and define ψ in a similar way with g . Then φ and ψ are G -transformations of E with no fixed points and $\varphi^2 = \psi^2 = \text{id}_E$.

Consider the group $H = \langle \varphi, \psi \rangle$ of G -transformations and its cyclic subgroup $N = \langle \psi\varphi \rangle$. Since both φ and ψ have order two, a simple computation will show that N is normal in H . We note that $\varphi N = \psi N$. Let E/N denote the orbit space. If $N.x \in E/N$ is an orbit, then $\varphi N.x = \psi N.x$. Hence φ and ψ induce the same transformation on E/N , which we denote by ρ .

The transformation ρ has order two and no fixed points. For suppose that $\rho(N.x) = N.x$ for some $x \in E$. Then $\varphi(\psi\varphi)^m.x = (\psi\varphi)^n.x$ for some $m, n \in \mathbb{Z}$. Hence $(\psi\varphi)^{-n}\varphi(\psi\varphi)^m.x = x$, which can be reduced by using

$$(\psi\varphi)^{-n}\varphi(\psi\varphi)^m = (\varphi\psi)^n\varphi(\psi\varphi)^m = \varphi(\psi\varphi)^{m+n}$$

to $\varphi(\psi\varphi)^{m+n}.x = x$.

2.1 A cancellation law

If $n + m$ is even, then $\varphi(\psi\varphi)^{m+n}$ and φ are conjugates, i.e. there is a $\delta \in \langle \varphi, \psi \rangle$, such that $\delta^{-1}\varphi\delta = \varphi(\psi\varphi)^{m+n}$. Simply choose $\delta = (\psi\varphi)^{\frac{m+n}{2}}$. We now get a contradiction because $\varphi(\psi\varphi)^{m+n}.x = x$ implies $\varphi.(\delta.x) = \delta.x$, but φ had no fixed points, and now $\delta.x$ is a fixed point. If $m + n$ is odd, then we get the same contradiction with ψ instead by choosing $\delta = \varphi(\psi\varphi)^{\frac{m+n-1}{2}}$, $\psi\delta.x = \delta.x$ for an appropriate $\delta \in \langle \varphi, \psi \rangle$.

Now, ρ partitions E/N in orbits. If Z consists of one element from each orbit (note the use of the axiom of choice), then $E/N = Z \coprod \rho(Z)$. If $\kappa : E \rightarrow E/N$ is the quotient map that maps an element to its orbit, then $E' = \kappa^{-1}(Z)$ and $E'' = \kappa^{-1}(\rho(Z))$ partitions E . Also, E' and E'' are interchanged by φ and also by ψ .

Let $\alpha : A \rightarrow E'$ be

$$\alpha(x) = \begin{cases} x, & x \in E' \cap A \\ \varphi(x), & x \in E'' \cap A \end{cases}$$

We claim that α is a G -transformation. Since φ is a G -transformation, and $\alpha(x) \in \{x, \varphi(x)\}$ for every $x \in A$, all that remains is to check that α is a bijection.

Injectivity: If $x \in E' \cap A$ and $y \in E'' \cap A$, then $\alpha(x) = x \in A$ while $\alpha(y) = \varphi(y) \in B$, so $\alpha(x) \neq \alpha(y)$, since A and B are disjoint. The other cases are trivial.

Surjectivity: Let $y \in E'$. If $y \in A$, then $y = \alpha(y)$. Otherwise $y \in B$, so $y = \varphi(x)$ for some $x \in A$ (actually $x = f^{-1}(y)$), and since φ interchanges E' and E'' , we must have $x \in E''$. We conclude $y = \varphi(x) = \alpha(x)$.

We have now shown that $A \sim_G E'$. In a similar way we can show $C \sim_G E'$, and the proof is complete, since then $A \sim_G C$. \square

Lemma 2.5 (Cancellation law). *Let $n \in \mathbb{N}$. If E_{2n} is G_{2n} -paradoxical, then E_n is G_n -paradoxical.*

Proof. Using the characterization of corollary 1.23 we may state the assumption as

$$E_{2n} = A \coprod B \quad \text{where } A \sim_{G_{2n}} E_{2n} \sim_{G_{2n}} B.$$

If we define $E' = E \times \{1, \dots, n\}$ and $E'' = E \times \{n+1, \dots, 2n\}$, then clearly

$$E_{2n} = E' \coprod E'' \quad \text{and} \quad E' \sim_{G_{2n}} E_n \sim_{G_{2n}} E''.$$

It follows from lemma 2.4 that $E_n \sim_{G_{2n}} E_{2n}$. This is equivalent to E_n being G_n -paradoxical. \square

Proof of proposition 2.3. In the case where n is a power of two, i.e. $n = 2^m$ for some $m \in \mathbb{N}$, the result is an easy consequence of the previous lemma using induction over m .

If n is not a power of two we may replace n by some larger integer which is a power of two and then derive the result for that integer. This is because if E_n is G_n -paradoxical, then E_m

is G_m -paradoxical for $m \geq n$. This can be proved by induction. We know that

$$E_n \preceq_{G_{2n+2}} E_{n+1} \preceq_{G_{2n+2}} E_{2n} \sim_{G_{2n+2}} E_n$$

so by the Banach-Schröder-Bernstein theorem $E_{n+1} \sim_{G_{2n+2}} E_n$, which shows that E_{n+1} is G_{2n+2} -paradoxical. Then E_{2n+2} must also be G_{2n+2} -paradoxical. By the previous lemma E_{n+1} is G_{n+1} -paradoxical. \square

2.2 Tarski's theorem

The proof of Tarski's theorem presented here is based on the Hahn-Banach theorem which plays a central role in functional analysis. It is an extension theorem used to extend continuous functionals defined on a subspace such that the extension is somehow well-behaved. There are several versions of the Hahn-Banach theorem. The one used here is stated in appendix B.

Let $L^\infty(E)$ be the real Banach space of bounded real functions on a set E with the supremum norm, and let $d^\infty(E)$ be the subspace generated by functions of the form $1_B - 1_A$ where $A, B \subseteq E$ and $A \sim_G B$. Here 1_A denotes the characteristic function of A . Further we denote by C the set of functions $f \in L^\infty(E)$ that satisfy $\inf\{f(x) \mid x \in E\} > 0$. Note that C is open.

We begin with a lemma using the Hahn-Banach theorem.

Lemma 2.6. *If $d^\infty(E) \cap C = \emptyset$, then there exists a finitely additive, G -invariant measure on X normalizing E .*

The converse is actually also true, but we will not prove it here. The idea is to look at the integral of functions in $L^\infty(E)$. But since we are dealing with finitely additive measures, things become complicated when we try to explain what we mean by an integral. We will return to this issue in chapter 3.

Proof. Suppose that $d^\infty(E) \cap C = \emptyset$. Then let V_0 be the subspace of $L^\infty(E)$ generated by $d^\infty(E)$ and 1_E , and let $F_0 : V_0 \rightarrow \mathbb{R}$ be

$$F_0(f + \lambda 1_E) = \lambda, \quad f \in d^\infty(E), \lambda \in \mathbb{R}.$$

It is easy to check that F_0 is a well-defined, linear functional on V_0 , and that $p : L^\infty(E) \rightarrow \mathbb{R}$

$$p(f) = \sup\{f(x) \mid x \in E\}$$

is a sublinear form on $L^\infty(E)$ which is an upper bound of F_0 . By the Hahn-Banach theorem it is possible to extend F_0 to a linear functional $F : L^\infty(E) \rightarrow \mathbb{R}$, such that $-p(-f) \leq$

2.2 Tarski's theorem

$F(f) \leq p(f)$ for all $f \in L^\infty(E)$. Now, define a function $\mu : \mathcal{P}(E) \rightarrow \mathbb{R}$ by

$$\mu(A) = F(1_A), \quad A \subseteq E.$$

First we will establish a few properties of μ , and then we will extend μ to $\mathcal{P}(X)$. For any $A \subseteq E$ we see

$$F(-1_A) \leq p(-1_A) \leq 0$$

so $F(1_A) \geq 0$, and hence μ takes only non-negative values. It is obvious that $\mu(\emptyset) = 0$ and $\mu(E) = 1$. If $B \subseteq E$ and $A \cap B = \emptyset$, then $1_{A \cup B} = 1_A + 1_B$, and since F is linear, μ is finitely additive. If $A \sim_G B$, then $1_B - 1_A \in d^\infty$, and so

$$\mu(B) - \mu(A) = F(1_B - 1_A) = 0,$$

which shows that

$$\mu(A) = \mu(B) \quad \text{if } A \sim_G B \text{ for } A, B \subseteq E \quad (2.1)$$

We shall now extend μ to all of $\mathcal{P}(X)$, so that the extension becomes a finitely additive, G -invariant measure on X . We define the extension $\nu : \mathcal{P}(X) \rightarrow [0, \infty]$ as follows: For $A \subseteq X$ it may happen that there is a partition of A into subsets $\{A_i\}_{i=1}^n$ where each $A_i \sim_G A'_i$ for some $A'_i \subseteq E$. In that case we define

$$\nu(A) = \sum_{i=1}^n \mu(A'_i).$$

Otherwise we define $\nu(A) = \infty$. We must check that ν is well-defined, and that ν is a finitely additive, G -invariant measure.

Well-defined: If

$$A = \prod_{i=1}^n A_i = \prod_{j=1}^m B_j, \quad \varphi_i(A_i) = A'_i, \quad \psi_j(B_j) = B'_j$$

where φ_i, ψ_j are G -transformations and $A'_i, B'_j \subseteq E$ for $i \in I_n, j \in I_m$, then define

$$A_{ij} = A_i \cap B_j, \quad A'_{ij} = \varphi_i(A_{ij}), \quad B'_{ij} = \psi_j(A_{ij}).$$

Then $A'_{ij} \sim_G A_{ij} \sim_G B'_{ij}$, so $\mu(A'_{ij}) = \mu(B'_{ij})$ by the invariance of μ . Since $\{A'_{ij}\}_{j=1}^m$ is a finite partition of A'_i and, similarly, $\{B'_{ij}\}_{i=1}^n$ partitions B'_j , we find

$$\sum_{i=1}^n \mu(A'_i) = \sum_{i=1}^n \sum_{j=1}^m \mu(A'_{ij}) = \sum_{j=1}^m \sum_{i=1}^n \mu(B'_{ij}) = \sum_{j=1}^m \mu(B'_j),$$

and so ν is well-defined.

Finite additivity: Suppose $A \cap B = \emptyset$ where $A, B \subseteq X$. If $\nu(A \cup B)$ is finite, then $\nu(A)$ and

$\nu(B)$ must both be finite. Hence if either $\nu(A)$ or $\nu(B)$ is infinite, then $\nu(A \cup B)$ is infinite, so $\nu(A \cup B) = \nu(A) + \nu(B)$.

If $\mu(A)$ and $\mu(B)$ are both finite, then it follows from the additivity of μ that $\nu(A \cup B) = \nu(A) + \nu(B)$. We leave the details to the reader.

G-invariance: If $A \sim_G B$, then A has finite measure if and only if B has finite measure. In the case where they both have finite measure the G -invariance follows from (2.1) and the way we defined our extension. Again, we leave the details to the reader. \square

We need yet another lemma.

Lemma 2.7. *Let $U_1, U_2 \subseteq E_n$, and let π_1 and π_2 be the canonical projection from E_n to E , but restricted to U_1 and U_2 , respectively. If $\pi_1^{-1}(x)$ and $\pi_2^{-1}(x)$ have the same cardinality for all $x \in E$, then there exists a G_n -transformation $\psi : U_1 \rightarrow U_2$ such that the diagram below commutes.*

$$\begin{array}{ccc} U_1 & \overset{\psi}{\dashrightarrow} & U_2 \\ \pi_1 \searrow & & \swarrow \pi_2 \\ & E & \end{array}$$

Proof. For any $x \in E$, put $F_x = \{q \in I_n \mid (x, q) \in U_1\}$ and $G_x = \{q \in I_n \mid (x, q) \in U_2\}$. By assumption F_x and G_x have the same cardinality, so there is a bijection $r_x : F_x \rightarrow G_x$, $r_x \in \mathcal{S}_n$. The set $P = \{r_x \in \mathcal{S}_n \mid x \in E\}$ is finite, simply because it is a subset of \mathcal{S}_n . Now put $\psi(x, q) = (x, r_x(q))$. Clearly $\psi : U_1 \rightarrow U_2$ is a bijection, and the diagram obviously commutes. Since P is finite, ψ is a G_n -transformation. \square

Theorem 2.8. *Let $E \subseteq X$. If there does not exist a finitely additive, G -invariant measure on X normalizing E , then E_n is G_n -paradoxical for some $n \in \mathbb{N}$.*

Proof. The idea of the following is to use lemma 2.6 to construct a G_{n+1} -transformation γ such that $\gamma(E_{n+1}) \subseteq E_n$. It will then follow rather easily that E_{n+1} is G_{n+1} -paradoxical.

By lemma 2.6 we can choose an element $f \in d^\infty(E) \cap C$ of the form

$$f = \sum_{i=1}^m \lambda_i (1_{B_i} - 1_{A_i})$$

where each $A_i \sim_G B_i$, and m is some natural number. Choose for each i a G -transformation $\gamma_i : A_i \rightarrow B_i$. Since C is open, there is an $r > 0$ such that

$$\{g \in L^\infty(E) \mid \|f - g\|_\infty < r\} \subseteq C.$$

For each λ_i it is possible to find a rational number q_i , such that $|\lambda_i - q_i| < \frac{r}{m}$. By the choice of q_i , the function

$$g = \sum_{i=1}^m q_i (1_{B_i} - 1_{A_i})$$

2.2 Tarski's theorem

is also in $d^\infty \cap C$. Further we may assume that every q_i is positive, because otherwise we could replace $1_{B_i} - 1_{A_i}$ by $1_{A_i} - 1_{B_i}$. By multiplying g by a sufficiently large natural number, we can assume that every q_i is an integer, and we can also assume that g is bounded below by 1. All in all we get

$$1 \leq g = \sum_{i=1}^n 1_{B_i} - 1_{A_i} \quad (2.2)$$

where some of the terms $1_{B_i} - 1_{A_i}$ may occur multiple times (in which case $n > m$). Adding $1_{A_i} + 1_{E \setminus A_i}$, $i = 1, \dots, n$ to (2.2) yields the inequalities

$$n + 1 \leq \sum_{i=1}^n 1_{B_i} + 1_{E \setminus A_i} \leq 2n.$$

Define

$$h = \sum_{i=1}^n 1_{B_i} + 1_{E \setminus A_i}$$

and put $U_2 = \{(x, q) \in E_{2n} \mid q \leq h(x)\}$. From the inequality $n + 1 \leq h$ we get $E_{n+1} \subseteq U_2$. Further define $\varphi : E_n \rightarrow E_{2n}$ by

$$\varphi(x, i) = \begin{cases} (\gamma_i(x), 2i - 1), & x \in A_i \\ (x, 2i), & x \in E \setminus A_i \end{cases}$$

We remark that $x \in A_i$ if and only if $\gamma_i(x) \in B_i$. It can easily be checked that φ is injective. If we denote the image of φ by U_1 then φ is a G_{2n} -transformation between E_n and U_1 .

Let π_1 and π_2 denote the projections from U_1 and U_2 to E . Now, by the construction of U_1 and U_2 we see that $|\pi_1^{-1}(x)| = h(x) = |\pi_2^{-1}(x)|$ for every $x \in E$. By lemma 2.7 there exists a G_{2n} -transformation ψ from U_1 to U_2 such that the diagram commutes.

$$\begin{array}{ccccc} E_n & \xrightarrow{\varphi} & U_1 & \xrightarrow{\psi} & U_2 \\ & & \searrow \pi_1 & & \swarrow \pi_2 \\ & & & & E \end{array}$$

Let γ be the restriction of $(\psi\varphi)^{-1}$ to E_{n+1} . Then γ is a G_{n+1} -transformation and $\gamma(E_{n+1}) \subseteq E_n$. Hence

$$E_{n+1} \supseteq E_n \supseteq \gamma(E_{n+1}) \supseteq \gamma(E_n) \supseteq \dots \supseteq \gamma^{n+1}(E_{n+1}) \supseteq \gamma^{n+1}(E_n),$$

so the sets $\gamma^i(E_{n+1} \setminus E_n)$, $i = 0, \dots, n$ and $\gamma^{n+1}(E_{n+1})$ are all pairwise disjoint. If we define $\rho_1 = \gamma^{n+1}$ and $\rho_2(x, i) = \gamma^{i-1}(x, n + 1)$, then ρ_1, ρ_2 are G_{n+1} -transformations with domain

E_{n+1} , and they have disjoint images. Hence

$$\rho_1(E_{n+1}) \sim_{G_{n+1}} E_{n+1} \sim_{G_{n+1}} \rho_2(E_{n+1}) \quad \text{and} \quad \rho_1(E_{n+1}) \cap \rho_2(E_{n+1}) = \emptyset,$$

so E_{n+1} is G_{n+1} -paradoxical. □

Now we have done all the preparation for proving Tarski's theorem.

Theorem 2.9 (Tarski's theorem). *Let G be a group acting on a set X , and let $E \subseteq X$. Then E is not G -paradoxical if and only if there exists a finitely additive, G -invariant measure on X normalizing E .*

Proof. If the measure exists, then the argument presented at the beginning of the chapter shows that E can not be G -paradoxical. Suppose, on the other hand, that no finitely additive, G -invariant measure normalizing E exists. The previous theorem shows that E_n is G_n -paradoxical for some $n \in \mathbb{N}$, and by proposition 2.3 we conclude that E is G -paradoxical. □

Combining Tarski's theorem with the strong form of the Banach-Tarski paradox we have the following remarkable result:

Corollary 2.10. *There does not exist a finitely additive, isometry-invariant measure on \mathbb{R}^3 normalizing a bounded subset with non-empty interior.*

Then what about \mathbb{R} and \mathbb{R}^2 ? It turns out, surprisingly, that such measures exist and they can even be chosen so that they extend the Lebesgue measure. That will be one of the main subjects of the next chapter. Vitali's construction of a non-measurable set shows, however, that such a measure can not be countably additive.

The purpose of this chapter is partly to prove that an analogy of the Banach-Tarski paradox does not exist in the plane or on the real line and partly to study groups bearing a certain kind of measure. The essential difference between the two-dimensional case and the three-dimensional case is that the isometry group \mathbb{E}_2 is solvable, but \mathbb{E}_3 is not. The essential link between the two subjects is Tarski's theorem, so in order to prove the absence of a paradox we construct a finitely additive measure normalizing the set, and the measure should be invariant under the group action.

3.1 Measures in groups

The way we constructed paradoxical decompositions in chapter 1 was by finding a paradoxical decomposition of the group acting on the set and then transferring the decomposition to the set. In much the same way our approach here will be to construct a measure on the group and somehow transfer the measure to a measure on the set upon which the group acts. This motivates the following definition.

Definition 3.1. If G is a group and $\mu : \mathcal{P}(G) \rightarrow [0, \infty]$ is finitely additive, left-invariant, i.e. $\mu(gA) = \mu(A)$ for all $g \in G$, $A \subseteq G$, and μ normalizes G , then μ is simply called a *measure* on G . A group for which such a measure exists is called *amenable*.

If μ is a measure on G and φ is an isomorphism from G to H , then $\mu \circ \varphi^{-1}$ is a measure on H , so if G and H are isomorphic, then either both are amenable or neither of them is. An application of Tarski's theorem gives the following characterization of amenable groups.

Corollary 3.2. *A group is amenable if and only if it is not paradoxical.*

As a corollary of proposition 1.8 we get:

Corollary 3.3. *If an amenable group G acts on a set X then X is not G -paradoxical.*

Proof. In the light of the preceding corollary this is just the contrapositive of proposition 1.8. □

Notice that this corollary only expresses the absence of paradoxicality of the whole set X and not of subsets $E \subseteq X$.

Let $L^\infty(G)$ denote the set of bounded real-valued functions on G . Recall from example 1.2 that G acts on $L^\infty(G)$ by letting $(g.f)(h) = f(g^{-1}h)$, where $g, h \in G$, $f \in L^\infty(G)$. When G is an amenable group it is possible to construct an integral defined on $L^\infty(G)$ taking real values and satisfying:

- (1) the integral is linear,
- (2) $\int f \, d\mu \geq 0$, if $f(g) \geq 0$ for all $g \in G$,
- (3) $\int 1_H \, d\mu = \mu(H)$ for every $H \subseteq G$,
- (4) $\int g.f \, d\mu = \int f \, d\mu$ for every $g \in G$.

For reasons of continuity the construction of the integral is omitted here but can be found in appendix C.

3.2 Some amenable groups

The definition of an amenable group may seem intangible, but we will now show that a wide class of well-known groups are amenable.

Proposition 3.4. *Every finite group is amenable.*

Proof. The counting measure on a finite group is a measure with the desired properties. \square

Proposition 3.5. *A subgroup of an amenable group is amenable.*

Proof. This is a consequence of corollary 1.7 and corollary 3.2, but a direct proof which does not rely on the difficult theorem of Tarski is actually quite simple: Suppose H is a subgroup of G , and μ is a measure on G . Let M be a set consisting of exactly one representative from each right coset with respect to H . Define

$$\nu(A) = \mu(A.M) = \mu\left(\bigcup_{m \in M} Am\right), \quad A \subseteq H.$$

It is easy to see that ν is finitely additive, H -invariant and $\nu(H) = 1$. We only show the invariance property and leave the rest for the reader to check.

Let $h \in H$ and $A \subseteq H$. Then

$$\nu(hA) = \mu\left(\bigcup_{m \in M} hAm\right) = \mu\left(h \bigcup_{m \in M} Am\right) = \mu\left(\bigcup_{m \in M} Am\right) = \nu(A).$$

\square

3.2 Some amenable groups

Proposition 3.6. *If N is a normal subgroup of an amenable group G , then the quotient G/N is amenable.*

Proof. Choose a measure μ on G . Each $A \subseteq G/N$ can be written $A = \{hN \mid h \in H\}$ for some subset $H \subseteq G$. We define $\nu : \mathcal{P}(G/N) \rightarrow [0, 1]$ by

$$\nu(A) = \mu \left(\bigcup_{h \in H} hN \right).$$

This is seen to be well-defined, i.e. not dependent of the choice of H in the representation of A as $\{hN \mid h \in H\}$. The desired properties of ν follow from those of μ . We will only demonstrate the invariance property. Suppose $gN \in G/N$ and $H \subseteq G$. Since N is normal

$$gN \cdot \{hN \mid h \in H\} = \{gNhN \mid h \in H\} = \{ghN \mid h \in H\}$$

showing that

$$\nu(gN \cdot A) = \mu \left(\bigcup_{h \in H} ghN \right) = \mu \left(g \cdot \bigcup_{h \in H} hN \right) = \mu \left(\bigcup_{h \in H} hN \right) = \nu(A),$$

where $A = \{hN \mid h \in H\}$.

□

The next theorems are not as straightforward to prove.

Proposition 3.7. *If N is a normal subgroup of G and both N and the quotient G/N are amenable, then G is amenable.*

Proof. Suppose μ_1, μ_2 are measures on N and G/N , respectively. For any $A \subseteq G$ define $f_A : G \rightarrow \mathbb{R}$ by $f_A(g) = \mu_1(N \cap g^{-1}A)$. If g_1 and g_2 are in the same coset with respect to N , say $g_2 = g_1n$, where $n \in N$, then

$$\mu_1(N \cap g_2^{-1}A) = \mu_1(N \cap (g_1n)^{-1}A) = \mu_1(N \cap n^{-1}g_1^{-1}A) = \mu_1(n^{-1}(N \cap g_1^{-1}A)) = \mu_1(N \cap g_1^{-1}A)$$

showing $f_A(g_2) = f_A(g_1)$. Therefore $F_A : G/N \rightarrow \mathbb{R}$ given by $F_A(gN) = f_A(g)$ is well-defined. Since $0 \leq \mu_1 \leq 1$, F is bounded. We may now define a measure ν on G by

$$\nu(A) = \int F_A \, d\mu_2.$$

We must check that ν is in fact a measure on G . Since $F_G(g) = 1_G$, $\nu(G) = 1$. Finite additivity of ν follows from the fact that $F_{A \cup B} = F_A + F_B$ if $A \cap B = \emptyset$. A simple computation will show that $f_{gA}(h) = f_A(g^{-1}h)$, and thus $F_{gA} = gN \cdot F_A$, and the invariance of the integral defined by μ_2 now gives $\nu(gA) = \nu(A)$. □

Corollary 3.8. *If G and H are amenable groups, then $G \times H$ is amenable, and in general any finite product of amenable groups is amenable.*

Proof. The latter follows from the first by induction. To prove amenability of $G \times H$, let $G^* = \{(g, 1) \mid g \in G\}$. Then $G^* \simeq G$, and so G^* is amenable, and G^* is normal in $G \times H$. Since $(G \times H)/G^* \simeq H$, the amenability of $G \times H$ follows from the preceding proposition. \square

The next theorem requires some acquaintance with topology and is based on Tychonoff's theorem: any product of compact spaces is compact in the product topology. We will not prove Tychonoff's theorem here. For a proof see any standard textbook on general topology.

A directed set is a partially ordered set (D, \leq) such that for any two elements $\alpha, \beta \in D$ there exists an element $\gamma \in D$ such that $\alpha \leq \gamma$ and $\beta \leq \gamma$. If D consists of subgroups of a certain group G , then the natural ordering is the subgroup relation. A directed set of subgroups of a group G is then a set D such that for any two subgroups G_α, G_β in D there exists a subgroup G_γ in D such that $G_\alpha \subseteq G_\gamma$ and $G_\beta \subseteq G_\gamma$.

Proposition 3.9. *If G is the union of a directed set of amenable subgroups, then G is amenable.*

Proof. Suppose G is the union of the directed set of subgroups $\{G_\alpha \mid \alpha \in I\}$ and each G_α is amenable. For each $\alpha \in I$ let μ_α be a finitely additive, G_α -invariant measure on G_α with $\mu_\alpha(G_\alpha) = 1$.

Consider the space $[0, 1]^{\mathcal{P}(G)}$ which by Tychonoff's theorem is compact in the product topology. For $\alpha \in I$ we denote by \mathcal{M}_α the set of those finitely additive $\mu : \mathcal{P}(G) \rightarrow [0, 1]$ such that $\mu(G) = 1$ and μ is G_α -invariant, i.e. $\mu(gA) = \mu(A)$ for every $g \in G_\alpha$ and $A \subseteq G$. By defining $\mu(A) = \mu_\alpha(A \cap G_\alpha)$ we see that $\mu \in \mathcal{M}_\alpha$, so each \mathcal{M}_α is non-empty. We will argue that \mathcal{M}_α is closed in $[0, 1]^{\mathcal{P}(G)}$. For $A \subseteq G$ we let $\pi_A : [0, 1]^{\mathcal{P}(G)} \rightarrow [0, 1]$ denote the projection map, i.e. $\pi_A(\mu) = \mu(A)$, which is continuous.

We proceed by showing that the complement of \mathcal{M}_α is open. Suppose $\mu \in [0, 1]^{\mathcal{P}(G)} \setminus \mathcal{M}_\alpha$. Then at least of one the following is true:

$$\mu(G) \neq 1, \quad \mu(A \cup B) \neq \mu(A) + \mu(B), \quad \mu(gA) \neq \mu(A)$$

for some $A, B \subseteq G$, $g \in G_\alpha$, where $A \cap B = \emptyset$.

If $\mu(G) \neq 1$, there is an open set $O \subseteq [0, 1]$ such that $\mu(G) \in O$ and $1 \notin O$; take for example $O = [0, 1[$. Then $\pi_G^{-1}(O)$ is open, and $\mu \in \pi_G^{-1}(O)$. By the choice of O , $\pi_G^{-1}(O) \cap \mathcal{M}_\alpha = \emptyset$.

If $\mu(A \cup B) \neq \mu(A) + \mu(B)$, choose $\varepsilon > 0$ such that $|\mu(A \cup B) - (\mu(A) + \mu(B))| > \varepsilon$. Then

3.2 Some amenable groups

the sets

$$O_1 = \{x \in [0, 1] \mid |x - \mu(A \cup B)| < \varepsilon/3\},$$

$$O_2 = \{x \in [0, 1] \mid |x - \mu(A)| < \varepsilon/3\},$$

$$O_3 = \{x \in [0, 1] \mid |x - \mu(B)| < \varepsilon/3\}$$

are open in $[0, 1]$, and hence $\pi_{A \cup B}^{-1}(O_1) \cap \pi_A^{-1}(O_2) \cap \pi_B^{-1}(O_3)$ is open in $[0, 1]^{\mathcal{P}(G)}$, contains μ and is contained in the complement of \mathcal{M}_α .

In much the same way we can handle the last case. If $\mu(gA) \neq \mu(A)$ then choose $\varepsilon > 0$ such that $|\mu(gA) - \mu(A)| > \varepsilon$. If

$$O_1 = \{x \in [0, 1] \mid |x - \mu(gA)| < \varepsilon/2\},$$

$$O_2 = \{x \in [0, 1] \mid |x - \mu(A)| < \varepsilon/2\},$$

then $\pi_{gA}^{-1}(O_1) \cap \pi_A^{-1}(O_2)$ is open, contains μ and is contained in the complement of \mathcal{M}_α .

Now that we have shown that each \mathcal{M}_α is non-empty and closed, we can finish the argument. If G_α and G_β are subgroups of G_γ then $\mathcal{M}_\gamma \subseteq \mathcal{M}_\alpha \cap \mathcal{M}_\beta$, so in particular $\mathcal{M}_\alpha \cap \mathcal{M}_\beta$ is non-empty. Hence $\{\mathcal{M}_\alpha \mid \alpha \in I\}$ has the finite intersection property. By compactness $\bigcap \{\mathcal{M}_\alpha \mid \alpha \in I\}$ is non-empty, and any μ in this set is a finitely additive, G -invariant measure, showing the amenability of G . \square

Before showing that any abelian group is amenable, we turn to one specific group, the integers $(\mathbb{Z}, +)$, and show that this group is amenable. The idea is inspired by the finite case, so we try to approximate the density of the set $A \subseteq \mathbb{Z}$ by measuring the density of finite sections. We use the notation ℓ^∞ for the real Banach space $\ell^\infty(\mathbb{N})$ with supremum norm.

Theorem 3.10. *The integers $(\mathbb{Z}, +)$ form an amenable group.*

Proof. Define $W_n = \{k \in \mathbb{Z} \mid -n \leq k \leq n\}$. Let $A \subseteq \mathbb{Z}$, and consider for each $n \in \mathbb{N}$ the number

$$f_n(A) = \frac{|A \cap W_n|}{|W_n|}.$$

What we would like to do is to define the measure of A to be the limit of $f_n(A)$ as n tends to infinity, but what if the limit does not always exist? An application of the Hahn-Banach theorem will get us around the problem.

Let c denote the subspace of ℓ^∞ consisting of the convergent sequences. The limit function $L_0 : c \rightarrow \mathbb{R}$ defined by

$$L_0(x) = \lim_{n \rightarrow \infty} x_n, \quad \text{where } x = (x_n) \in c,$$

is a linear functional on c , which is bounded above by the sublinear form $p(x) = \sup x_n$. By the Hahn-Banach theorem there is an extension of L_0 to all of ℓ^∞ , call it L , such that $-p(-x) \leq L(x) \leq p(x)$ for every $x \in \ell^\infty$.

Note that $0 \leq f_n(A) \leq 1$, so the sequence $(f_n(A))_{n \in \mathbb{N}}$ belongs to ℓ^∞ . We may now define $\mu : \mathcal{P}(\mathbb{Z}) \rightarrow \mathbb{R}$ by

$$\mu(A) = L((f_n(A))_{n \in \mathbb{N}}), \quad \text{for all } A \subseteq \mathbb{Z}.$$

We must check that μ is the desired measure. For $x \in \ell^\infty$ where $x_n \geq 0$ for all $n \in \mathbb{N}$ we have $-L(x) \leq p(-x) \leq 0$, so $L(x) \geq 0$. Hence $\mu(A) \geq 0$ for any $A \subseteq \mathbb{Z}$.

Clearly $\mu(\mathbb{Z}) = 1$, and the additivity of μ follows from the additivity of L and the fact that $f_n(A \cup B) = f_n(A) + f_n(B)$ if $A \cap B = \emptyset$. Since the sets $A \cap W_n$ and $(A + k) \cap W_n$ has a difference in cardinality of at most $2k$, we get

$$|f_n(A) - f_n(A + k)| \leq \frac{2k}{2n + 1} \rightarrow 0 \quad \text{for } n \rightarrow \infty.$$

so $\mu(A + k) = \mu(A)$, i.e. μ is translation invariant. □

We can now show that any abelian group is amenable.

Theorem 3.11. *Every abelian group is amenable.*

Proof. Any group is the union of its finitely generated subgroups, and the finitely generated subgroups form a directed set of subgroups. Hence by theorem 3.9 it suffices to prove that a finitely generated abelian group is amenable. By the fundamental theorem of finitely generated abelian groups such a group is isomorphic to a product of a finite group and \mathbb{Z}^n , where $n \in \mathbb{N} \cup \{0\}$ is the rank of the group. We have already shown that finite groups and \mathbb{Z} are amenable. Now, a reference to corollary 3.8 completes the proof. □

Corollary 3.12. *Every solvable group is amenable.*

Proof. This is a consequence of the preceding theorem combined with proposition 3.7 using induction on the length of a subnormal series witnessing the solvability. □

3.3 The real line and the plane

The isometry groups of \mathbb{R} and \mathbb{R}^2 are solvable (see appendix for a proof), and hence by corollary 3.12 they are amenable. By corollary 3.3, \mathbb{R}^n is not \mathbb{E}_n -paradoxical if $n = 1, 2$. This is, of course, interesting in itself, but we are more interested in showing that no interval or square is paradoxical. To prove this we will construct an isometry-invariant extension of the Lebesgue measure defined on *all* subsets of \mathbb{R}^n . As mentioned at the end of the last chapter, Vitali's example shows that such a measure can not be countably additive, but all we need is

3.3 The real line and the plane

finite additivity, and this turns out to be obtainable. We will proceed via the Hahn-Banach theorem.

Theorem 3.13. *If G is an amenable group of isometries of \mathbb{R}^n , then there exists a finitely additive, G -invariant measure defined on all subsets of \mathbb{R}^n , which extends the Lebesgue measure.*

Proof. Let m denote the Lebesgue measure on \mathbb{R}^n . Let $\mathcal{L} = V_0$ denote the vector space of measurable, integrable functions $f : \mathbb{R}^n \rightarrow \mathbb{R}$, and let V be the space of functions $f : \mathbb{R}^n \rightarrow \mathbb{R}$ such that $|f| \leq f_0$ for some $f_0 \in \mathcal{L}$. Clearly V is a real vector space with \mathcal{L} as a subspace. Any group G of isometries acts on V by $(g.f)(x) = f(g^{-1}.x)$, where $g \in G$, $f \in V$, $x \in \mathbb{R}^n$. The action is linear

$$g.(\lambda_1 f_1 + \lambda_2 f_2) = \lambda_1 g.f_1 + \lambda_2 g.f_2.$$

It is well-known that \mathcal{L} is stable under the action of G , i.e. $g.f \in \mathcal{L}$ if $f \in \mathcal{L}$, so G also acts on \mathcal{L} .

Define a linear functional F_0 on \mathcal{L} by

$$F_0(f_0) = \int f_0 \, dm, \quad f_0 \in \mathcal{L}.$$

That is, F_0 is integration with respect to Lebesgue measure. Since the Lebesgue integral is invariant under isometries, F_0 is G -invariant.

If $f \in V$ and $f_0 \in \mathcal{L}$ satisfy $|f| \leq f_0$, then for any $f_1 \in \mathcal{L}$ such that $f \leq f_1$, we must have $-f_0 \leq f \leq f_1$ so

$$F_0(-f_0) = \int -f_0 \, dm \leq \int f_1 \, dm = F_0(f_1).$$

Hence

$$\inf\{F_0(f_1) \mid f_1 \in \mathcal{L} \text{ and } f \leq f_1\}$$

is a real number not less than $F_0(-f_0)$, and we may define a sublinear form p on V by

$$p(f) = \inf\{F_0(h) \mid h \in \mathcal{L} \text{ and } f \leq h\}.$$

It is easy to check that $p(\lambda f) = \lambda p(f)$ for $\lambda \geq 0$. To check $p(f + g) \leq p(f) + p(g)$, note that if $f \leq f_1$ and $g \leq g_1$, then $f + g \leq f_1 + g_1$. This shows the inequality in the following:

$$\begin{aligned} p(f + g) &= \inf\{F_0(h_1) \mid h_1 \in \mathcal{L} \text{ and } f + g \leq h_1\} \\ &\leq \inf\{F_0(f_1 + g_1) \mid f_1, g_1 \in \mathcal{L} \text{ and } f \leq f_1, g \leq g_1\} \\ &= \inf\{F_0(f_1) \mid f_1 \in \mathcal{L} \text{ and } f \leq f_1\} + \inf\{F_0(g_1) \mid g_1 \in \mathcal{L} \text{ and } g \leq g_1\} \\ &= p(f) + p(g) \end{aligned}$$

What is also important, and trivial to see, is that p is G -invariant, meaning $p(g.f) = p(f)$.

Now, by the Hahn-Banach theorem there exists a linear functional $F : V \rightarrow \mathbb{R}$ that extends F_0 , also bounded by p . But F need not be G -invariant.

For any $f \in V$ define $\varphi_f : G \rightarrow \mathbb{R}$ by $\varphi_f(g) = F(g^{-1}.f)$. Since

$$F(g^{-1}.f) \leq p(g^{-1}.f) = p(f),$$

and

$$F(g^{-1}.f) \geq -p(-(g^{-1}.f)) = -p(g^{-1}.(-f)) = -p(-f),$$

we see that φ_f is bounded above by $p(f)$ and below by $-p(-f)$. In particular $\varphi_f \in L^\infty(G)$, so choosing a measure ν on G we may define a function $\mu : \mathcal{P}(\mathbb{R}^n) \rightarrow \mathbb{R} \cup \{\infty\}$ by

$$\mu(A) = \int \varphi_{1_A} d\nu, \quad \text{if } 1_A \in V,$$

and $\mu(A) = \infty$ if $1_A \notin V$. Here again 1_A denotes the characteristic function of A . We claim that μ is a finitely additive, G -invariant measure, and that μ extends m .

μ is an extension of m : Let A be a Lebesgue measurable subset of \mathbb{R}^n . If $1_A \in V_0$, then

$$\varphi_{1_A}(g) = F(g^{-1}.1_A) = F_0(g^{-1}.1_A) = F_0(1_A) = m(A),$$

and so φ_{1_A} is the constant function with value $m(A)$. Using that $\nu(G) = 1$ we get $\mu(A) = m(A)$. If $1_A \notin V_0$, then $1_A \notin V$ and so $\mu(A) = \infty = m(A)$.

The following observation will be useful. Using the linearity of F and of $g^{-1} \in G$ when acting on elements in V a simple computation will show that

$$\varphi_{\lambda_1 f_1 + \lambda_2 f_2}(g) = \lambda_1 \varphi_{f_1}(g) + \lambda_2 \varphi_{f_2}(g) \tag{3.1}$$

for $f_1, f_2 \in V, \lambda_1, \lambda_2 \in \mathbb{R}, g \in G$. If $f \in V$ and $f(x) \geq 0$ for every $x \in \mathbb{R}^n$, then since $\varphi_{-f} \leq p(-f)$,

$$\int \varphi_{-f} d\nu \leq \int p(-f) d\nu = p(-f) \cdot 1 \leq 0,$$

and from this and (3.1),

$$\int \varphi_f d\nu = - \int \varphi_{-f} d\nu \geq 0$$

so in particular $\mu(A) \geq 0$ for every $A \subseteq \mathbb{R}^n$.

If $A \cap B = \emptyset$ for $A, B \subseteq \mathbb{R}^n$, then $1_{A \cup B} = 1_A + 1_B$. This combined with (3.1) shows that μ is finitely additive.

If $g, h \in G$ and $f \in V$, then

$$\varphi_{g.f}(h) = F(h^{-1}(g.f)) = F((g^{-1}h)^{-1}.f) = \varphi_f(g^{-1}h) = (g.\varphi_f)(h),$$

3.3 The real line and the plane

so the G -invariance of ν and the fact that $1_{gA} = g.1_A$ yield that

$$\mu(gA) = \int \varphi_{1_{gA}} \, d\nu = \int \varphi_{g.1_A} \, d\nu = \int g.\varphi_{1_A} \, d\nu = \int \varphi_{1_A} \, d\nu = \mu(A).$$

This completes the proof, since μ is the desired measure. \square

Corollary 3.14. *The Lebesgue on \mathbb{R} and \mathbb{R}^2 , respectively, has a finitely additive, isometry-invariant extension of to all sets.*

Proof. It is a consequence of the previous theorem since \mathbb{E} and \mathbb{E}_2 are solvable (cf. appendix A) and hence amenable. \square

Corollary 3.15. *If G is an amenable group of isometries of \mathbb{R}^n , no bounded subset of \mathbb{R}^n with non-empty interior is G -paradoxical. In particular, no bounded subset of \mathbb{R} or \mathbb{R}^2 with non-empty interior is paradoxical using isometries.*

Proof. Suppose A is a bounded subset of \mathbb{R}^n with non-empty interior, and let μ be a finitely additive, isometry-invariant extension of the Lebesgue measure defined on all subsets of \mathbb{R}^n . Since A has non-empty interior, $\mu(A) > 0$ and the boundedness of A ensures $\mu(A) < \infty$. A paradoxical decomposition of A would imply $\mu(A) = 2\mu(A)$, but this is impossible. \square

Corollary 3.15 shows that no interval in \mathbb{R} is \mathbb{E} -paradoxical, and no square in \mathbb{R}^2 is \mathbb{E}_2 -paradoxical. Other paradoxical decompositions in \mathbb{R}^2 do exist, however. But these either rely on groups that do not preserve distance (but they may still preserve area), use sets with empty interior or use unbounded sets.

The isometry groups of the Euclidean spaces

To every metric space M is associated the group of isometries consisting of the distance-preserving bijections from M to itself. In the case of $M = \mathbb{R}^n$ we know exactly how these bijections look. The isometry group of \mathbb{R}^n is denoted by \mathbb{E}_n and contains the subgroups T_n of translations and O_n of orthogonal mappings. Any element $f \in \mathbb{E}_n$ can be written in the form $f = t \circ A$ where t is a translation and A is an orthogonal map. Moreover, t and A are uniquely determined by f . The map $f \mapsto A$ is thus a well-defined homomorphism from \mathbb{E}_n onto O_n with kernel T_n , so $T_n \triangleleft \mathbb{E}_n$ and by the isomorphism theorem, $\mathbb{E}_n/T_n \simeq O_n$.

Each element of $A \in O_n$ has a representation as an orthogonal $n \times n$ -matrix with determinant ± 1 . The map $A \mapsto \det A$ is a homomorphism from O_n onto $\{\pm 1\}$, and the kernel is denoted SO_n . The matrix representing an element of SO_n has determinant 1, and SO_n is called the special orthogonal group.

The composed map $f \mapsto \det A$ is a homomorphism, and the kernel is denoted $S\mathbb{E}_n$. Hence $S\mathbb{E}_n$ is normal in \mathbb{E}_n , and by the isomorphism theorem the quotient $\mathbb{E}_n/S\mathbb{E}_n$ is isomorphic to $\{\pm 1\}$, the cyclic group of order 2. T_n is contained in $S\mathbb{E}_n$, and since T_n is normal in \mathbb{E}_n it is also normal in $S\mathbb{E}_n$.

In the thesis we will need the following essential theorem about the Euclidean isometry groups.

Theorem A.1. \mathbb{E}_1 and \mathbb{E}_2 are solvable, and \mathbb{E}_n is not solvable for $n \geq 3$.

Proof. Here we will only prove the first half of the theorem. That \mathbb{E}_n is not solvable for $n \geq 3$ is actually a consequence of lemma 1.11, since \mathbb{E}_n contains a free subgroup of rank 2, but we will not go into the details.

We have the following subnormal series

$$\{\text{id}\} \triangleleft T_n \triangleleft S\mathbb{E}_n \triangleleft \mathbb{E}_n$$

The factor groups are

1. $\mathbb{E}_n/S\mathbb{E}_n \simeq \{\pm 1\}$, which is cyclic of order two, hence abelian,

2. $\mathbb{SE}_n/T_n \simeq \mathbb{SO}_n$, which is abelian in the case $n = 1, 2$,

3. $T_n/\{\text{id}\} \simeq T_n$ which is abelian.

Since all the factor groups are abelian in the case $n = 1, 2$, we conclude that \mathbb{E}_1 and \mathbb{E}_2 are solvable. \square

The crucial observation in the proof above is that \mathbb{SO}_n is abelian when $n = 1, 2$.

The Hahn-Banach Theorem

The Hahn-Banach theorem is used several times throughout the thesis, so we will mention it here for the sake of completeness. For more details and a proof consult for example [Rud].

Definition B.1. A real function $p : V \rightarrow \mathbb{R}$ on a real vector space V is called a *sublinear form* if $p(x + y) \leq p(x) + p(y)$ and $p(\lambda x) = \lambda p(x)$ for each $x, y \in V$ and $\lambda \geq 0$.

Example B.2. An important example which is used several times in the thesis is the case where V is the set $L^\infty(X)$ of bounded real functions on a set X . If we define $p(f) = \sup\{f(x) \mid x \in X\}$, for $f \in L^\infty(X)$, then

$$p(f + g) = \sup_{x \in X} (f(x) + g(x)) \leq \sup_{x \in X} f(x) + \sup_{x \in X} g(x) = p(f) + p(g)$$

and for $\alpha \geq 0$,

$$p(\alpha f) = \sup_{x \in X} \alpha f(x) = \alpha \sup_{x \in X} f(x) = \alpha p(f).$$

Hence p is a sublinear form on $L^\infty(X)$.

There are several versions of the Hahn-Banach theorem. We are content with the following.

Theorem B.3 (The Hahn-Banach theorem). *Suppose V_0 is a subspace of real vector space V and $p : V \rightarrow \mathbb{R}$ is a sublinear form on V . If $F_0 : V_0 \rightarrow \mathbb{R}$ is linear and $F_0(x) \leq p(x)$ on V_0 , then there exists a linear functional $F : V \rightarrow \mathbb{R}$ such that*

$$F(x) = F_0(x) \quad \text{for } x \in V_0$$

and

$$-p(-x) \leq F(x) \leq p(x) \quad \text{for } x \in V.$$

 Integration with finitely additive probability measures

Pre-integrals

Let G be a group and let \mathcal{A} be an algebra (not necessarily a σ -algebra) on G . Suppose μ is a finitely additive measure on G with $\mu(G) = 1$. In the thesis we only consider the case $\mathcal{A} = \mathcal{P}(G)$, in which case things become much simpler, because we have no concerns about measurability (and G -stability – see below).

Let $\ell^\infty = \ell^\infty(G, \mathcal{A})$ denote the Banach space of measurable, bounded functions $f : G \rightarrow \mathbb{R}$ with norm $\|f\|_\infty = \sup\{|f(g)| \mid g \in G\}$. Note that here we equip \mathbb{R} with the algebra generated by the open sets. We consider the class of simple functions $\mathcal{S} \subseteq \ell^\infty$ which attain only finitely many values (negative values are allowed). We will need the following lemma.

Lemma C.1. *\mathcal{S} is dense in ℓ^∞ .*

Proof. Let $f \in \ell^\infty$ and $\varepsilon > 0$ be given. Since f is bounded, $f(G)$ is contained in a compact interval I . There exists a finite covering of I by measurable subsets A_1, \dots, A_N of \mathbb{R} with $\text{diam}A_i < \varepsilon$. We can assume, that $A_i \cap A_j = \emptyset$ for $i \neq j$. Put $B_i = A_i \cap f(G)$. We can also assume that each B_i is non-empty. Now, define $G_i = f^{-1}(B_i) = f^{-1}(A_i)$. Every G_i is measurable, because $A_i \in \mathcal{A}$, and f is measurable, so by choosing $b_i \in B_i$ the function

$$s = \sum_{i=1}^N b_i 1_{G_i}$$

is simple. $\{G_i\}_{i=1}^N$ partitions G and for $g \in G_i$ both $f(g) \in B_i$ and $s(g) \in B_i$, so since

$$\text{diam}B_i \leq \text{diam}A_i < \varepsilon$$

we see that

$$\|s - f\|_\infty \leq \max_{1 \leq i \leq N} \{\text{diam}B_i\} < \varepsilon.$$

as desired. □

For the simple functions we can define the pre-integral $I : \mathcal{S} \rightarrow \mathbb{R}$ by

$$I\left(\sum_{i=1}^N b_i 1_{G_i}\right) = \sum_{i=1}^N b_i \mu(G_i).$$

We note that this is a finite number because μ is a probability measure. The representation of the simple function can be shown to be immaterial, and so I is well-defined. Clearly I satisfies the well-known rules

$$I(s + t) = I(s) + I(t) \quad I(as) = aI(s)$$

for $s, t \in \mathcal{S}$, $a \in \mathbb{R}$ and also $I(s) \geq 0$ if $s(g) \geq 0$ for every $g \in G$.

Integrals

The pre-integral $I : \mathcal{S} \rightarrow \mathbb{R}$ satisfies $|I(s)| \leq \|s\|_\infty$, i.e. I is a contraction. Because of this I is uniformly continuous. We are now in need of the following proposition

Proposition C.2 (Extension theorem). *Let M_0 be a dense subset of a metric space M and let Y be a complete metric space. If $f_0 : M_0 \rightarrow Y$ is uniformly continuous, then there exists a (unique) continuous extension $f : M \rightarrow Y$ of f_0 .*

Proof. For any $m \in M$ let $m_n \rightarrow m$ where each $m_n \in M_0$. Since (m_n) is a Cauchy sequence and f_0 is uniformly continuous, $(f_0(m_n))$ is a Cauchy sequence in Y . By completeness the sequence converges to a $y \in Y$. Define $f(m) = y$. Note that y is independent of the choice of (m_n) , so f is well-defined. From the construction it follows that f is continuous. This proves the existence. Uniqueness is obvious. \square

Take $M = \ell^\infty$. The combination of lemma C.1 with the previous proposition allows us to define the integral of a function $f \in \ell^\infty$

$$\int f \, d\mu$$

through the unique extension of the pre-integrals. Additivity and the like will follow from continuity. We note that since characteristic functions are simple we especially get the pleasant rule

$$\int 1_A \, d\mu = \mu(A)$$

for every $A \in \mathcal{A}$.

Translation invariance

A group G acts on itself by left translation. If $gA \in \mathcal{A}$ for every $g \in G$ and $A \in \mathcal{A}$, we say that \mathcal{A} is G -stable, and the action of G on itself extends to one on \mathcal{A} . G also acts on ℓ^∞ by

$$(g.f)(x) = f(g^{-1}.x), \quad g, x \in G, f \in \ell^\infty.$$

Suppose that \mathcal{A} is G -stable and that the measure μ is G -invariant, that is $\mu(gA) = \mu(A)$ for each $g \in G$, $A \in \mathcal{A}$. Then the integral becomes G -invariant, that is

$$\int g.f \, d\mu = \int f \, d\mu, \quad g \in G, f \in \ell^\infty.$$

We first prove this for characteristic functions, then simple functions, and the result is then a consequence of lemma C.1. For $A \in \mathcal{A}$ we have $g.1_A = 1_{gA}$, so

$$\int g.1_A \, d\mu = \int 1_{gA} \, d\mu = \mu(gA) = \mu(A) = \int 1_A \, d\mu.$$

For a simple function the result is a consequence of the linearity of I . Now, let $f \in \ell^\infty$ be any function. By lemma C.1 there is a sequence $\{s_n\}$ of simple functions that converges to f . Each s_n has an integral which is invariant under the action. Since the integral of f is the limit of $I(s_n)$ as n tends to infinity the result follows.

To sum up we know that whenever G is a group with a G -stable algebra \mathcal{A} that has a finitely additive, G -invariant measure μ , it is possible to define integration on ℓ^∞ with respect to μ , and this integral will respect addition, scalar multiplication and positivity, i.e. if f is non-negative then also the integral of f is non-negative. Further the integral of the characteristic function of a measurable subset A gives the measure of A , and the integral is invariant under the action of G on ℓ^∞ .

Bibliography

- [C-SGH] T. Ceccherini-Silberstein, R. Grigorchuk and P. de la Harpe, *Amenability and paradoxical decompositions for pseudogroups and for discrete metric spaces*, Proc. Steklov Inst. Math. 224 (1999), 57-97.
- [ES] Elek, Gábor, and Sós, Vera T., *Paradoxical Decompositions and Growth Conditions*, Combinatorics, Probability and Computing (2005) 14, 81-105.
- [HS] De la Harpe, P., and G. Skandalis, *Un résultat de Tarski sur les actions moyennables de groupes et les partitions paradoxales*, l'Enseignement Mathématique, Vol. 32 (1986), 121-138.
- [Mun] Munkres, James R., *Topology*, Prentice Hall, Second edition 2000. First published in 1975.
- [Rud] Rudin, Walter, *Functional Analysis*, McGraw-Hill, 1973.
- [Wag] Wagon, Stan, *The Banach-Tarski paradox*, Cambridge University Press, Cambridge, 1994. First published in 1985.