The Hahn-Banach Theorem implies the existence of a non-Lebesgue measurable set

Matthew Foreman, Friedrich Wehrung

THE HAHN-BANACH THEOREM IMPLIES THE EXISTENCE OF A NON LEBESGUE-MEASURABLE SET

Matthew Foreman and Friedrich Wehrung
Ohio State University (U.S.A.)
CNRS, Université de Caen. France.

§0. Introduction.

Few methods are known to construct non Lebesgue-measurable sets of reals: most standard ones start from a well-ordering of $\mathbb{R}$, or from the existence of a non-trivial ultrafilter over ω, and thus need the axiom of choice AC or at least the Boolean Prime Ideal theorem BPI (see [5]). In this paper we present a new way for proving the existence of non-measurable sets using a convenient operation of a discrete group on the Euclidian sphere. The only choice assumption used in this construction is the Hahn-Banach theorem, a weaker hypothesis than BPI (see [9]). Our construction proves that the Hahn-Banach theorem implies the existence of a non-measurable set of reals. This answers questions in [9], [10]. (Since we do not even use the countable axiom of choice, we cannot assume the countable additivity of Lebesgue measure; e.g. the real numbers could be a countable union of countable sets.)

In fact we prove (under Hahn-Banach theorem) that there is no finitely additive, rotation invariant extension of Lebesgue measure to $\mathcal{P}(\mathbb{R}^3)$. Notice that Hahn-Banach implies the existence of a finitely additive, isometry invariant extension of Lebesgue measure to $\mathcal{P}(\mathbb{R}^2)$ (see [14]).

We use standard set-theoretical notation and terminology. For example, if X is any set, $\mathcal{P}(X)$ is the power set of X. If $A \subseteq X$ and $f : X \to Y$ is a map, then $f[A]$ is the image of A under f. Furthermore, ω is the set of all natural numbers.

We assume ZF throughout this paper; no choice assumption (even countable) is made.

§1. Definitions.

First, let us give one of the many equivalent statements of the Hahn-Banach theorem. We use the version [11]:

The Hahn-Banach Theorem. *Let E be a vector space over the reals, let S be a subspace of E, and f be a linear functional on S. Let p be a map $E \to \mathbb{R}$ such that whenever $x, y \in E$ and $\lambda \geq 0$, we have $p(\lambda x) = \lambda p(x)$ and $p(x+y) \leq p(x) + p(y)$. Then there is a linear functional $\bar{f}$ on E, extending f, such that $(\forall x \in E)(\bar{f}(x) \leq p(x))$.*

Definition. *If B is a Boolean algebra, a finitely additive probability measure on B (from now on a measure) is a map $\mu : B \to [0,1]$ such that $\mu(1_B) = 1$ and $\mu(x \vee y) = \mu(x) + \mu(y)$ whenever $x \wedge y = 0$.*

It is known that $ZF+$ Hahn-Banach implies that every Boolean algebra has a measure (actually in ZF without choice, this last statement is equivalent to the Hahn-Banach theorem, see [7,15]). It also yields the following statement for collections of Boolean algebras:

Proposition 1. *(ZF+Hahn-Banach theorem) Let $\langle B_i : i \in I\rangle$ be a sequence of Boolean algebras (with I not necessarily well-orderable). Then there exists $\langle \mu_i : i \in I\rangle$ such that for each $i \in I$, μ_i is a measure on B_i.*

Proof. Let $(B, e_i)_{i\in I}$ be the direct sum of $(B_i)_{i\in I}$ in the category of Boolean algebras: so, for every $i \in I$, e_i is an homomorphism $B_i \to B$ (elements of B are formal Boolean combinations of elements of the B_i with no other relations than those from the B_i; one can prove that e_i is one-to-one). By the Hahn-Banach theorem there is a measure μ on B. Put $\mu_i = \mu \circ e_i$. ■

Definition. *A universally measured space is an ordered pair (Ω, μ) where Ω is a set and μ is a measure on the Boolean algebra $\mathcal{P}(\Omega)$. A group G is said to act by measure preserving transformations on (Ω, μ) when G acts on Ω and $\mu(gA) = \mu(A)$ for all $g \in G$ and $A \in \mathcal{P}(\Omega)$.*

We are going to be mainly concerned about the following measure existence statement:

Definition. *Let a group G act on a set Ω. $IM(\Omega, G)$ is the statement "there is a G-invariant measure on $\mathcal{P}(\Omega)$".*

In the case of a group acting on itself, we get the following classical definition.

Definition. *A group G is amenable when there is a measure μ on $\mathcal{P}(G)$ such that $\mu(Ag) = \mu(A)$ for all $g \in G$, $A \in \mathcal{P}(G)$.*

Assuming the Hahn-Banach theorem many groups are amenable, including finite groups, solvable groups and their extensions. The best known non-amenable group is the free group on two generators.

Proposition 2. *(Classical) [14] - The free group on two generators, F_2, is not amenable.*

For all integers $n \geq 1$, denote by O_n the isometry group of S^{n-1} (with Euclidian norm), $SO_n = \{u \in O_n : \det(u) = +1\}$, where $S^n = \{x \in \mathbb{R}^{n+1} : \|x\| = 1\}$ is the n-dimensional Euclidian sphere. One can prove in ZFC that $IM(S^n, SO_{n+1})$ does not hold for $n \geq 2$, and thus SO_{n+1} is not amenable (see [14]). On the other hand, in [10] and [13], the authors construct models of $ZF + DC$ in which $IM(S^n, O_{n+1})$ holds for every $n \geq 1$ (in [13], the measure is just normalized Lebesgue-measure).

A group G acts on a set Ω *freely* when for all $g \in G$, $x \in \Omega$, $gx = x$ implies $g = 1$.

§2. The main results.

We start with a classical result.

Proposition 3. *Assume $IM(S^2, SO_3)$. Then there is a free measure-preserving action of F_2 on some universally measured space (Ω, μ).*

Proof. Consider a subgroup of SO_3 isomorphic to F_2, [14] and D the subset of S^2 consisting of the union of all the possible orbits of fixed points of elements of $F_2\backslash\{1\}$. D is countable since each orbit is effectively countable and it is easy to distinguish fixed points of elements of F_2 acting on S^2. Hence D is the image of a function with domain $\{0, 1\} \times F_2 \times F_2$. (Recall, we do not know that a countable union of countable sets is countable.) Let μ be the witness to $IM(S^2, SO_3)$. Since F_2 acts freely on $S^2\backslash D$, we will be done if we can show $\mu(D) = 0$.

In [14] it is shown that every SO_3-invariant finitely additive measure on S^2 gives each countable set measure zero. We paraphrase the proof given there and check that it works without AC.

It clearly suffices to find a rotation g such that for all $k \in \omega\backslash\{0\}$, $g^k D \cap D = \emptyset$. Since then $\{g^k D : k \in \omega\}$ is an infinite collection of pairwise disjoint subsets of S^2 of the same μ-measure. Let $\langle a_n : n \in \omega\rangle$ be an enumeration of D. Let ℓ be a line through the origin missing D. Let $A_n = \{g \in SO(3) : g$ is a rotation about ℓ and for some $i \neq j \in \omega$, $g^n a_i = a_j\}$. Then A_n is countable in a canonical way, since each $g \in A_n$ is determined by a_i and a_j. Hence $\cup A_n$ is countable. Choose a rotation g about ℓ such that $g \notin \cup A_n$ and g has infinite order. Then for all $n \geq 1$, $g^n D \cap D = \emptyset$. ■

Another example is with $IM({}^\omega 2, G)$ where ${}^\omega 2$ is the Cantor space with its canonical metric and G its group isometries (see [12]).

Our main theorem is:

Theorem 4. *(ZF+Hahn-Banach) - Let a group G act freely and measure-preserving on a universally measured space (Ω, μ). Then G is amenable.*

Proof. (Note the similarity to [6].)

Denote by Ω/G the set of orbits of Ω modulo G.

By Proposition 1, there is a sequence $\langle \mu_{[x]} : [x] \in \Omega/G\rangle$ such that for each $[x] \in \Omega/G$, $\mu_{[x]}$ is a measure on $\mathcal{P}([x])$. For each $A \subseteq G$, let $a : \Omega \to [0,1]$ be the following function: $a(x) = \mu_{[x]}(Ax)$; define $\lambda : \mathcal{P}(G) \to [0,1]$ by $\lambda(A) = \int a(x)\, d\mu(x)$. Note that $x \mapsto a(x)$ is a measurable function since (Ω, μ) is a universally measured space; the integration here is essentially Lebesgue integration, and it does not appeal to any choice (no limit theorems are needed).

We claim that λ is a measure on $\mathcal{P}(G)$, invariant under right translaton.

Note that $\lambda(G) = 1$. If A, B are two disjoint subsets of G and a, b, c are the functions corresponding to A, B, $A \cup B$ respectively, then $(\forall x \in \Omega)(c(x) = a(x) + b(x))$. Hence $\lambda(A \cup B) = \lambda(A) + \lambda(B)$.

Finally, if $B = Ag$ for some $g \in G$ and a, b are the functions corresponding to A and B then, for all $x \in \Omega$,

$$\begin{aligned} b(x) &= \mu_{[x]}(Bx) = \mu_{[x]}(Ag\,x) \\ &= \mu_{[x]}(A(gx)) = \mu_{[gx]}(A(gx)) = a(gx)\,. \end{aligned}$$

Hence $\lambda(B) = \int b(x)\, d\mu(x) = \int a(gx)\, d\mu(x) = \int a(x)\, d\mu(x) = \lambda(A)$ since g is μ-measure preserving. ■

Corollary 1. - *ZF+Hahn-Banach implies not $IM(S^2, SO_3)$. Thus, there is a non-Lebesgue measurable subset of S^2.*

Proof. Propositions 2, 3 and Theorem 4. ■

Note that in the last part of the statement above, S^2 could be replaced by many other spaces, like $\mathbb{R}^n$, $n \geq 1$. (See §3 for details).

Corollary 2. *If H is generic for the partial ordering adding ω_1 random reals to a model V of ZFC and $V(\mathbb{R})$ is the smallest model of set theory containing V and reals of $V[H]$, then $V(\mathbb{R})$ does not satisfy the Hahn-Banach theorem.*

Proof. $V(\mathbb{R})$ is the model considered considered by D. Pincus and R. Solovay in [10]. It satisfies $IM(S^n, SO_{n+1})$ for all $n \geq 1$, and thus $IM(S^2, SO_3)$; we conclude by Corollary 1. ∎

Another way to see Corollary 1 is the following:

Corollary 3. *If F_2 acts freely on $\Omega = S^2 \backslash D$ (D as in the proof of Proposition 3) by rotations, and if $\langle \mu_{[x]} : [x] \in \Omega/F_2 \rangle$ is any assignment of finitely additive probability measures $\mu_{[x]}$ on $\mathcal{P}([x])$, then there are $A \subseteq F_2$ and $\alpha \in [0,1]$ such that $\{x : \mu_{[x]}(Ax) < \alpha\}$ is not Lebesgue measurable. Further the set A can be isolated explicitly (see [14]).* ∎

§3. Appendix. Lebesgue measure without countable choice.

Ordinarily, the theory of Lebesgue measure is developed with use of AC_ω. The use of AC_ω allows one to use arbitrary Borel sets. In this section we explore how to use "coded" Borel sets to eliminate the necessity of AC_ω in many applications. For example, we would still like the existence of non-measurable set to be independent from the reference space (here, S^2). The aim of this section is to show how to adapt the proofs of the "classical" theory (with AC_ω) to the study of Lebesgue-measure in a totally choiceless context. The ideas here date from [13].

In order to get as many measurable sets as possible, the classical outer measure construction (see [4]) seems convenient enough. This construction, which we will sketch in $\mathbb{R}$, works as well in $\mathbb{R}^n$ or in much more abstract spaces.

Define the outer measure of $A \subseteq \mathbb{R}$ by the greatest lower bound of all sums $\sum\limits_{n\in\omega} \text{length}(I_n)$ where I_n are intervals, and $A \subseteq \bigcup\limits_{n\in\omega} I_n$; call it $\mu^*(A)$. Say that A is Lebesgue-measurable when for all $X \subseteq \mathbb{R}$, $\mu^*(X) = \mu^*(X \cap A) + \mu^*(X \backslash A)$. Note $\mathcal{M} = \{A \subseteq \mathbb{R};\ A \text{ is Lebesgue-measurable}\}$, $\mu = \mu^* \mid \mathcal{M}$. It is still possible to prove that $\mathcal{M}$ is a Boolean subalgebra of $\mathcal{P}(\mathbb{R})$ and that μ is a finitely additive function $\mathcal{M} \to [0, \infty]$, and that $\mathcal{M}$ contains all open sets. But one cannot prove any more that $\mathcal{M}$ is a σ-algebra (since $\mathbb{R}$ can be a countable union of countable sets, see [5]). So, instead of considering Borel subsets of $\mathbb{R}$, consider those which have a *code*, as e.g. in [12]; a Borel code is essentially a real, encoding the "construction" of some Borel set. Similarly, say that $(A_n)_{n\in\omega}$ is coded sequence of Borel sets when there is a sequence $(c_n)_{n\in\omega}$ such that for every n, c_n is a code for A_n. And then, we can prove the following properties of $(\mu, \mathcal{M})$:

(a) *$\mathcal{M}$ is Boolean subalgebra of $\mathcal{P}(\mathbb{R})$, containing all coded Borel subsets of $\mathbb{R}$.*

(b) *μ is a finitely additive map $\mathcal{M} \to [0, \infty]$, and whenever $(A_n)_{n\in\omega}$ is a disjoint coded sequence of Borel sets, we have:*

$$\mu\left(\bigcup_{n\in\omega} A_n\right) = \sum_{n\in\omega} \mu(A_n)\,.$$

(c) *A subset* $A \subseteq \mathbb{R}$ *is in* $\mathcal{M}$ *iff for all* $\varepsilon > 0$ *and all coded Borel* B *with* $\mu(B) < \infty$, *there are coded Borel* F *and* U *such that* $F \subseteq A \cap B \subseteq U$ *and* $\mu(U \backslash F) < \varepsilon$.

(Actually, it is enough to check when B is a bounded interval, and U can be chosen as an open set, F as a closed set.)

(d) μ *is* σ*-finite: there is a coded sequence* $(A_n)_{n\in\omega}$ *of Borel sets such that* $\mathbb{R} = \bigcup_{n\in\omega} A_n$ *and* $(\forall n \in \omega)(\mu(A_n) < \infty)$. *(Take* $A_n = [-n, n]$*.)*

The precautions needed by elimination of AC_ω in the classical proof of (a) and (d) above (see [4]) make the proof somewhat more lengthy, but without real difficulties. Note that in (c), the assumption $\mu(B) < \infty$ does not seem to be removable without countable choice.

Let us call the μ above the Lebesgue measure on $\mathbb{R}$; a similar construction yields Lebesgue measure on $\mathbb{R}^n$, for all $n \geq 1$.

More generally, let us set the following definition:

Definition. *A coded Borel space is an ordered pair* $(\Omega, \mathcal{B})$ *where* Ω *is a coded Borel subset of the Hilbert cube* ${}^\omega[0,1]$ *and* $\mathcal{B}$ *is the algebra of coded Borel subsets of* Ω.

We can naturally extend this definition by taking all isomorphic images; this way, all usual spaces of analysis - like $\mathbb{R}^n$, S^n, or ${}^\omega 2$, together with their coded Borel subsets, become coded Borel spaces. Anyway, even without using countable choice, it turns out that the following is true:

Proposition 5. *Let* $(\Omega, \mathcal{B})$ *be an uncountable coded Borel space. Then there is a coded Borel isomorphism from* $(\Omega, \mathcal{B})$ *onto* $(I, \mathcal{B}_I)$*, where* $I = [0, 1]$ *and* $\mathcal{B}_I$ *is the algebra of coded Borel subsets of* I. ∎

Here, a coded Borel isomorphism $(\Omega, \mathcal{B}) \to (I, \mathcal{B}_I)$ is naturally a bijection $f : \Omega \to I$ such that the neighborhood diagrams of f and f^{-1} are coded Borel.

Now, let us give the new definition of measure we are going to use:

Definition. *Let* $(\Omega, \mathcal{B})$ *be a coded Borel space. A regular measure on* $(\Omega, \mathcal{B})$ *is a map* $\mu : \mathcal{M} \to [0, \infty]$ *such that* $(\mu, \mathcal{M})$ *satisfies conditions (a) to (d) above, with* Ω *instead of* $\mathbb{R}$. *Say that* μ *is nonatomic when* $(\forall x \in \Omega)(\mu(\{x\}) = 0)$.

The essential isomorphism theorem between these measure spaces is still valid (after a suitable reformation). It can be stated the following way:

Proposition 6. *Let* μ *be a regular, nonatomic measure on a coded Borel space* $(\Omega, \mathcal{B})$, *with* $\mu(\Omega) = 1$. *Then there are* $N \subseteq \Omega$, $D \subseteq [0, 1]$ *and* $f : \Omega \to [0, 1]$ *such that, if* ℓ *is Lebesgue measure on* $[0, 1]$,

(i) $N \in \mathcal{B}$, D *is countable,* $\mu(N) = \ell(D) = 0$.

(ii) f *is a coded Borel isomorphism* $\Omega \backslash N \to [0, 1] \backslash D$.

(iii) For all B *in* $\mathcal{B}$, $f[B]$ *is coded Borel in* $[0, 1]$ *and* $\mu(B) = \ell(f[B])$.

Outline of Proof (See [11]). First, notice that by (b) and $\mu(\Omega) = 1$, Ω is uncountable. So, by proposition 5, without loss of generality, $\Omega = [0, 1]$ and $\mathcal{B}$ is the algebra of coded Borel subsets of $[0, 1]$. Then, define $f : [0, 1] \to [0, 1]$ by $f(x) = \mu([0, x])$. Then, D is just

$\{y \in [0,1] : f^{-1}\{y\}$ has nonempty interior$\}$ and N is $f^{-1}[D]$. (iii) is proven by induction on a code of B, and it uses nonatomicity of μ. ∎

Now, Proposition 6 has an immediate corollary:

Corollary 1. *Let μ be a regular, nonatomic measure on a coded Borel space $(\Omega, \mathcal{B})$, with $\mu(\Omega) \neq 0$. Then the following are equivalent:*

(i) Every subset of Ω is μ-measurable.

(ii) Every subset of $[0,1]$ is Lebesgue-measurable. ∎

(To prove (i) $\Rightarrow$ (ii), one has to use σ-finiteness, nonatomicity of μ and $\mu(\Omega) \neq 0$; for (ii) $\Rightarrow$ (i), use characterisation (c) above of μ-measurability).

In particular, every subset of $\mathbb{R}^n$ ($n \geq 1$) is Lebesgue-measurable iff every subset of $[0,1]$ is Lebesgue-measurable (which is well-known in the classical theory using countable choice). Let LM be the latter statement.

Now, define Lebesgue measure v_n on S^n as being the image under $x \mapsto \frac{x}{\|x\|}$ of Lebesgue measure on $B^{n+1}\backslash\{0\}$, where B^{n+1} is the Euclidian closed ball of $\mathbb{R}^{n+1}$ of volume 1.

Corollary 2. *LM implies $IM(S^n, SO_{n+1})$ for all $n \geq 1$.*

Proof. If LM holds, then v_n is defined on $\mathcal{P}(S^n)$ by the previous corollary; so v_n witnesses $IM(S^n, SO_{n+1})$. ∎

More precisely, the result would be the same with a rotation-invariant extension of Lebesgue-measure on $\mathcal{P}(S^2)$; thus, the results of the previous paragraph imply for example that *Hahn-Banach theorem implies nonexistence of a rotation-invariant extension of Lebesgue-measure to a (finitely additive) measure on $\mathcal{P}(\mathbb{R}^3)$.*

Further notes. Theorem 4 could be formulated as follows: *"If G is a nonamenable group acting freely on a set Ω and if μ is a G-invariant finitely additive probability measure defined on a G-invariant subalgebra of $P(\Omega)$, then Ω has non-measurable subsets (w.r.t. μ)".* Now, while this paper was printed, the second author showed, under the same hypotheses, that in the G-equidecomposability type semigroup of Ω (see [14]), $n[\Omega] = (n+1)[\Omega]$ for some integer n, effectively computable from the number of pieces necessary to a paradoxical decomposition of G. For the action of F_2 described above, we can get $n = 5$, which is somewhat disappointing since it is not known whether the cancellation law (see [14]) follows from HB (it follows from BPI). But independently, J. Pawlikowski proved using ideas from this paper, that one can actually take $n = 1$, that is, $[\Omega] = 2[\Omega]$; thus, HB implies the Banach-Tarski paradox. See [8] for more details.

BIBLIOGRAPHY

[1] M. Foreman, "Amenable groups acting on the natural numbers, an independence result". Bulletin AMS, October 1989, vol. 21, N° 2.

[2] K. Gödel, "The consistency of the axiom of choice and the generalized continuum hypothesis with the axioms of set theory". Ann. of Math. Studies N° 3. Princeton University Press, 1940.

[3] F. Greenleaf, "Invariant means on Topological Groups". Van Nostrand, Reinhold Company, 1969.

[4] P. R. Halmos, "Measure Theory". Van Nostrand, Rienhold Company, 1950.

[5] T. Jech, "The axiom of choice". North-Holland Publishing Company.

[6] A. Kechris, "Amenable equivalence relations and Turing degrees". Handwritten notes, 1988.

[7] W. A. J. Luxemburg, "Reduced products of the real number system". Applications of model theory to algebra, analysis and probability. Holt, Reinhart and Winston, 1969.

[8] J. Pawlikowski, "The Hahn-Banach Theorem implies the Banach-Tarski Paradox". To appear.

[9] D. Pincus, "The strength of Hahn-Banach's Theorem". Victoria symposium on non-standard analysis. Springer lecture notes n° 369, 1974 (pp. 203-248).

[10] D. Pincus and R. Solovay, "Definability of measures and ultrafilters". Journal of Symbolic Logic 42, n° 2, 1977, (pp. 179-190).

[11] H. L. Royden, "Real Analysis". The Macmillan Compnay, New York, 1968.

[12] T. Slaman and J. Steel, "Definable functions on Degrees". Cabal seminar 1981-85, appeared in 1988, Springer Verlag lecture notes.

[13] R. Solovay, "A model of set theory in which every set of reals in Lebesgue-measurable". Annals of Math. 1970, (pp. 1-56).

[14] S. Wagon, "The Banach-Tarski Paradox". Cambridge University Press, 1984.

[15] J. Conway, "A course in Functional Analysis". Springer-Verlag, 1985.