

Chapter 10

The Meaning of Means

FINITELY ADDITIVE MEASURES, EXTENSIONS, AND AMENABLE GROUPS

Overview. In the last chapter we used the Markov–Kakutani Theorem to produce for every compact abelian group G an “invariant mean”: a mean in the algebraic dual space of $B(G)$ fixed by the adjoint of every translation map on $B(G)$. The Riesz Representation Theorem then provided a G -invariant, regular, Borel probability measure to represent this mean via integration on $C(G)$. Thus was born Haar measure for G .

In this chapter we will scale back the role of topology and observe that for each abelian group G this “Markov–Kakutani” mean easily provides a G -invariant, finitely additive “probability measure” on $\mathcal{P}(G)$, the collection of all subsets of G . We’ll examine the significance of such set functions. In the compact case, might one of them extend Haar measure? Which non-abelian groups support such “measures”? Such questions will lead (next chapter) into the study of “paradoxical decompositions,” most notably the celebrated Banach–Tarski Paradox.

Prerequisites. A little: measure theory, group theory, functional analysis.

10.1 Means and Finitely Additive Measures

We’ve previously attached to a set S the following cast of characters:

- $\mathcal{P}(S)$: The collection of all subsets of S .
- $B(S)$: The vector space of all bounded, real-valued functions on S .
- $B(S)^\sharp$: The *algebraic dual* of $B(S)$; all the linear functionals on $B(S)$.
- $\mathcal{M}(S)$: The *means* on $B(S)$; the collection of positive linear functionals Λ on $B(S)$ “normalized” so that $\Lambda(1) = 1$. We’ve noted that $\mathcal{M}(S)$ is a nonempty, convex subset of $B(S)^\sharp$ (Exercise 9.13).
- $\omega(S)$: The *weak-star topology* on $B(S)^\sharp$; the restriction to $B(S)^\sharp$ of the product topology of $\mathbb{R}^{B(S)}$. We’ve seen that $\mathcal{M}(S)$ is $\omega(S)$ -compact (Corollary 9.17).

“Measures” from Means. Each mean Λ on $B(S)$ naturally defines a function μ on the collection $\mathcal{P}(S)$ of all subsets of S :

$$\mu(E) = \Lambda(\chi_E) \quad (E \in \mathcal{P}(S)), \quad (10.1)$$

where χ_E denotes the characteristic function of E ($\equiv 1$ on E and $\equiv 0$ off E).

Exercise 10.1. For μ as defined above, show that:

- (a) $\mu(S) = 1$.
- (b) μ is *monotone*: $E \subset F \subset S \implies \mu(E) \leq \mu(F)$.
- (c) $\mu(E) \leq 1$ for every $E \subset S$.

The linearity of Λ translates into *finite additivity* for μ : if $\{E_1, E_2, \dots, E_n\}$ is a finite, pairwise disjoint collection of subsets of S then $\chi_{\cup_k E_k} = \sum_k \chi_{E_k}$, so

$$\mu\left(\bigcup_k E_k\right) = \Lambda(\chi_{\cup_k E_k}) = \Lambda\left(\sum_k \chi_{E_k}\right) = \sum_k \Lambda(\chi_{E_k}) = \sum_k \mu(E_k).$$

Definition 10.1. A *finitely additive probability measure* on $\mathcal{P}(S)$ is a finitely additive function $\mu : \mathcal{P}(S) \rightarrow [0, 1]$ with $\mu(S) = 1$.

In this terminology the argument above established:

Proposition 10.2. Each mean Λ on $B(S)$ induces via Eq. (10.1) a *finitely additive probability measure* μ on $\mathcal{P}(S)$.

The exercise below shows that conversely each finitely additive probability measure μ on $\mathcal{P}(S)$ gives rise to a mean on $B(S)$, created as a sort of “Riemann integral.”

Exercise 10.2 (Means from “Measures”). Let $\mathcal{S}(S)$ denote the collection of “simple functions” on S , i.e., the functions $f : S \rightarrow \mathbb{R}$ that take on only finitely many values.

Given a finitely additive probability measure μ on $\mathcal{P}(S)$ and a simple function f on S with distinct values $\{a_j\}_{j=1}^n$, let $E_j = f^{-1}(a_j)$ and define $\Lambda(f) := \sum_{j=1}^n a_j \mu(E_j)$.

- (a) Check that \mathcal{S} is a vector space on which the functional Λ is positive and linear, and that Λ obeys the inequality promised for means by Exercise 9.14.
- (b) Show that Λ has a unique extension to a mean on $B(S)$ [Hint: Show that Λ is continuous if \mathcal{S} is given the “sup-norm” $\|\cdot\|$ defined on $B(S)$ by Eq. (9.10)].

Invariant Means. Theorem 9.19 told us that if Φ is a commutative family of self-maps of the set S , then $B(S)$ has a mean Λ that is Φ -invariant in the sense that $C_\varphi^\# \Lambda = \Lambda$ for every $\varphi \in \Phi$, where $C_\varphi : B(S) \rightarrow B(S)$ is the composition operator $f \rightarrow f \circ \varphi$ defined on p. 114. The finitely additive probability measure μ that Λ induces on $\mathcal{P}(S)$ via Eq. (10.1) inherits this Φ -invariance:

$$\mu(\varphi^{-1}(E)) = \Lambda(\chi_{\varphi^{-1}(E)}) = \Lambda(\chi_E \circ \varphi) = (C_\varphi^\# \Lambda)(\chi_E) = \Lambda(\chi_E) = \mu(E)$$

for every $E \subset S$ and $\varphi \in \Phi$. In summary:

Theorem 10.3. *If Φ is a commutative family of self-maps of a set S then there is a finitely additive probability measure μ on $\mathcal{P}(S)$ that is Φ -invariant in the sense that $\mu(\varphi^{-1}(E)) = \mu(E)$ for every $E \in \mathcal{P}(S)$ and every $\varphi \in \Phi$.*

Corollary 10.4. *If G is a commutative group then there exists a finitely additive probability measure μ on $\mathcal{P}(G)$ that is G -invariant in the sense that $\mu(gE) = \mu(E)$ for each $g \in G$ and $E \in \mathcal{P}(G)$.*

Proof. Apply Theorem 10.3 with $S = G$ and Φ the collection of “translation maps” $x \rightarrow g^{-1}x$ for g and x in G . □

Suppose in Theorem 10.3 we take Φ to be the group of rotations of \mathbb{R}^2 about the origin and S to be either of the following subsets of \mathbb{R}^2 : the closed unit disc \mathbb{B}^2 , or its boundary \mathbb{T} , the unit circle. In either case Φ is a commutative family of self-maps of S , hence Theorem 10.3 yields:

Corollary 10.5. *Both $\mathcal{P}(\mathbb{B}^2)$ and $\mathcal{P}(\mathbb{T})$ support a rotation-invariant, finitely additive probability measure.*

The question arises for either case: can such a finitely additive, rotation-invariant probability measure be chosen to agree, on Borel sets, with normalized Lebesgue measure. Similarly, for every compact abelian group, must the invariant measure promised by Corollary 10.4 agree on Borel sets with Haar measure? We’ll study this matter of invariant extension in the next section. Not surprisingly, it will involve the Hahn–Banach Theorem.

10.2 Extending Haar Measure

Suppose G is a compact abelian group. We now know that G has both:

- a G -invariant regular probability measure μ on its Borel sets (Haar measure: Corollary 9.20), and
- a G -invariant *finitely additive* probability measure ν on its collection $\mathcal{P}(G)$ of *all* subsets (Theorem 10.3).

Since ν arose from a G -invariant mean on $B(G)$, and μ (via the Riesz Representation Theorem) from the restriction of that mean to $C(G)$, one might suspect that ν extends Haar measure from the Borel subsets of G to all of $\mathcal{P}(G)$, i.e., that the restriction of ν to the Borel subsets of G is μ . Surprisingly, this need not be the case; Banach proved in 1923 that it fails for the circle group \mathbb{T} .

There exists a rotation-invariant, finitely additive probability measure on $\mathcal{P}(\mathbb{T})$ that does not extend Haar measure [6, Théorème 20].

We’ll see later that for a compact group: there can be *at most one* Haar measure (Chap. 12), and that *there always is* a Haar measure (Chap. 13). In particular, for the unit circle \mathbb{T} , normalized Lebesgue measure is the unique rotation-invariant regular Borel measure. Thus Banach’s result can be restated:

There exists a rotation-invariant, finitely additive probability measure on $\mathcal{P}(\mathbb{T})$ whose restriction to the Borel subsets of \mathbb{T} is not countably additive.

In view of Banach's result, it makes sense to ask if Haar measure on G can be extended to a finitely additive G -invariant probability measure on $\mathcal{P}(G)$. Thanks to the Markov–Kakutani Theorem the answer is affirmative, with the desired extension of Haar measure following from an “invariant” version of the Hahn–Banach Theorem. First recall the usual version:

The Hahn–Banach Theorem. *Suppose V is a vector space over the real field and $p: V \rightarrow \mathbb{R}$ is a gauge function on V , i.e.,*

- (a) $p(u + v) \leq p(u) + p(v)$ for all $u, v \in V$, and
- (b) $p(av) = ap(v)$ for every $a \in \mathbb{R}$ with $a \geq 0$ and every $v \in V$.

Suppose W is a linear subspace of V and Λ is a linear functional on W for which $\Lambda(w) \leq p(w)$ for all $w \in W$. Then Λ has a linear extension $\tilde{\Lambda}$ to V such that

$$\tilde{\Lambda}(v) \leq p(v) \quad \text{for all } v \in V.$$

Now consider that problem of extending Haar measure μ from the Borel subsets of a compact abelian group G to a finitely additive measure ν on $\mathcal{P}(G)$. The measure μ induces, via integration, a G -invariant linear functional Λ on $C(G)$, where we now view $C(G)$ as a linear subspace of $B(G)$. In order to make the desired extension of μ it will be enough to extend Λ to a G -invariant mean on $B(G)$. This will be accomplished by:

Theorem 10.6 (The “Invariant” Hahn–Banach Theorem). *Suppose V is a vector space and \mathcal{G} is a commutative family of linear transformations $V \rightarrow V$. Suppose W is a linear subspace of V that is taken into itself by every transformation in \mathcal{G} , and that p is a gauge function on V that is “ \mathcal{G} -subinvariant” in the sense that*

$$p(\gamma(v)) \leq p(v) \quad \text{for every } v \in V \text{ and } \gamma \in \mathcal{G}.$$

If Λ is a \mathcal{G} -invariant linear functional on W that is dominated by p , i.e.,

$$\Lambda \circ \gamma = \Lambda \quad \text{for all } \gamma \in \mathcal{G} \quad \text{and} \quad \Lambda(v) \leq p(v) \quad \text{for all } v \in W,$$

then Λ has a \mathcal{G} -invariant linear extension to V that is dominated on V by p .

Proof. Endow V^\sharp , the algebraic dual of V , with the weak-star topology ω induced on it by V . Let \mathcal{K} be the collection of all linear extensions of Λ to V that are dominated on V by p . Clearly \mathcal{K} is a convex subset of V^\sharp . By the (usual) Hahn–Banach Theorem, \mathcal{K} is nonempty.

Claim: \mathcal{K} is weak-star compact in V^\sharp .

Proof of Claim. By Corollary 9.15 we need only show that \mathcal{K} is pointwise bounded on V and weak-star closed in V^\sharp . If $\tilde{\Lambda} \in \mathcal{K}$ then for every $x \in V$ we have, in addition to the defining property $\tilde{\Lambda}(x) \leq p(x)$, also $-\tilde{\Lambda}(x) = \tilde{\Lambda}(-x) \leq p(-x)$. Thus

$$-p(-x) \leq \tilde{\Lambda}(x) \leq p(x) \quad (x \in V, \tilde{\Lambda} \in \mathcal{K}) \tag{10.2}$$

so \mathcal{K} is pointwise bounded on V .

To see that \mathcal{K} is weak-star closed in V^\sharp , suppose $\Lambda_0 \in V^\sharp$ is a weak-star limit point of \mathcal{K} . We wish to show that $\Lambda_0 \in \mathcal{K}$, i.e., that Λ_0 is an extension of Λ from W to V that's dominated by p . To see that Λ_0 extends V , fix $w \in W$ and $\varepsilon > 0$. Then the weak-star basic neighborhood

$$N(\Lambda_0, \{w\}, \varepsilon) = \{\Lambda \in V^\sharp : |\Lambda(w) - \Lambda_0(w)| < \varepsilon\}$$

contains a linear functional $\Lambda_1 \in \mathcal{K}$. Thus $|\Lambda_0(w) - \Lambda(w)| = |\Lambda_0(w) - \Lambda_1(w)| < \varepsilon$, whereupon $\Lambda_0(w) = \Lambda(w)$ because ε is an arbitrary positive number; hence Λ_0 is an extension of Λ to V .

Similarly, fix $v \in V$ and $\varepsilon > 0$. Choose $\Lambda_2 \in \mathcal{K} \cap N(\Lambda_0, \{v\}, \varepsilon)$. Then $|\Lambda_0(v) - \Lambda_2(v)| < \varepsilon$, so $\Lambda_0(v) < \Lambda_2(v) + \varepsilon \leq p(v) + \varepsilon$, hence $\Lambda_0(v) \leq p(v)$, once again by the arbitrariness of ε . This completes the proof of the Claim.

Finally, since each $\gamma \in \mathcal{G}$ is a linear map $V \rightarrow V$, it has an *adjoint* $\gamma^\sharp : V^\sharp \rightarrow V^\sharp$. Let $\mathcal{G}^\sharp := \{\gamma^\sharp : \gamma \in \mathcal{G}\}$. One checks easily that \mathcal{G}^\sharp is a commutative family of linear maps on V^\sharp , each of which, thanks to the \mathcal{G} -subinvariance of the gauge function p , takes \mathcal{K} into itself. By Theorem 9.18 each map γ^\sharp is ω -continuous, hence the triple $(V^\sharp, \mathcal{K}, \mathcal{G}^\sharp)$, with V^\sharp carrying its weak-star topology, satisfies the hypotheses of the Markov–Kakutani theorem.

Conclusion: There exists $\tilde{\Lambda} \in K$ fixed by \mathcal{G}^\sharp , i.e.,

$$\tilde{\Lambda} \circ \gamma = \gamma^\sharp(\tilde{\Lambda}) = \tilde{\Lambda} \quad \text{for every } \gamma \in \mathcal{G}.$$

This functional $\tilde{\Lambda}$ is the desired \mathcal{G} -invariant extension of our original one Λ . □

Here, stated in generality, is our application to extension of invariant measures.

Corollary 10.7. *Let S be a compact Hausdorff space upon which acts a commutative family Φ of continuous mappings. Suppose μ is a (countably additive) Φ -invariant probability measure on the Borel subsets of S . Then μ extends to a Φ -invariant, finitely additive probability measure on $\mathcal{P}(S)$.*

Proof. Let Λ be the positive linear functional defined on $C(S)$ by integration against μ . By the invariance of μ and the change-of-variable formula of measure theory, Λ is invariant for each of the composition operators C_φ on $C(S)$ in the sense that $\Lambda \circ C_\varphi = \Lambda$ for each $\varphi \in \Phi$. Define the gauge function p on $B(S)$ by

$$p(f) = \|f\| = \sup_{s \in S} f(s) \quad (f \in B(S)).$$

Clearly: p is C_Φ -invariant (in the sense that $p \circ C_\varphi = p$ for every $\varphi \in \Phi$), and $\Lambda \leq p$ on $C(S)$.

The Invariant Hahn–Banach Theorem now supplies an extension of Λ to a linear functional $\tilde{\Lambda}$ on $B(S)$ that's also dominated by p , and is invariant for each mapping C_φ for $\varphi \in \Phi$. Upon applying inequality (10.2) to our gauge function p , we see that

$$\inf_{s \in S} f(s) \leq \tilde{\Lambda}(f) \leq \sup_{s \in S} f(s) \quad (f \in B(S)),$$

so if $f \geq 0$ on S then $\tilde{\Lambda}(f) \geq 0$, i.e., $\tilde{\Lambda}$ is a positive linear functional on $B(S)$. Since $\tilde{\Lambda}(1) = \Lambda(1) = 1$, the functional $\tilde{\Lambda}$ is a mean on $B(S)$. The desired extension $\tilde{\mu}$ of μ to $\mathcal{P}(S)$ now emerges from Eq. (10.1) with $\tilde{\Lambda}$ in place of Λ , the Φ -invariance of $\tilde{\mu}$ following from the C_Φ -invariance of $\tilde{\Lambda}$. \square

For our original problem of extending Haar measure on a compact abelian group G , we take in Corollary 10.7: $S = G$ and $\Phi =$ the set of translation maps $x \rightarrow g^{-1}x$ for g and x in G . The result:

Corollary 10.8. *For each compact abelian group G , Haar measure has an extension to a finitely additive G -invariant measure on $\mathcal{P}(G)$.*

Since the group of rotations of \mathbb{R}^2 about the origin is abelian, Corollary 10.7 yields

Corollary 10.9. *There is a rotation-invariant, finitely additive probability measure on the closed unit disc of \mathbb{R}^2 that extends Lebesgue area measure from the Borel sets to all subsets. The unit circle supports a similar extension of normalized arc-length measure.*

Our final application of the invariant Hahn–Banach theorem involves the creation of a notion of “limit” for every bounded real sequence. We’ll use the notation ℓ^∞ for the space of all such sequences.

Corollary 10.10 (Banach limits). *There exists a positive, translation-invariant linear functional Λ on ℓ^∞ such that*

$$\liminf_{n \rightarrow \infty} f(n) \leq \Lambda(f) \leq \limsup_{n \rightarrow \infty} f(n) \quad (f \in \ell^\infty).$$

Proof. Let c denote the space of real sequences $f: \mathbb{N} \rightarrow \mathbb{R}$ for which $\lambda(f) = \lim_{n \rightarrow \infty} f(n)$ exists (in \mathbb{R}). For $f \in \ell^\infty$ let

$$p(f) = \limsup_{n \rightarrow \infty} f(n).$$

Then p is a gauge function on ℓ^∞ , and $\lambda \leq p$ on c . For $k \in \mathbb{N}$ define the “translation map” T_k on ℓ^∞ by

$$T_k f(n) = f(n+k) \quad (f \in \ell^\infty, n \in \mathbb{N}).$$

Thus $\mathcal{T} = \{T_k : k \in \mathbb{N}\}$ is a commutative family of linear transformations $\ell^\infty \rightarrow \ell^\infty$ for each of which: the subspace c is taken into itself, and both λ and p are invariant. Thus the Invariant Hahn–Banach Theorem applies and produces a \mathcal{T} -invariant extension Λ of λ to ℓ^∞ with $\Lambda \leq p$ on ℓ^∞ . By inequality (10.2):

$$\liminf_{n \rightarrow \infty} f(n) = -p(-f) \leq \Lambda(f) \leq p(f) = \limsup_{n \rightarrow \infty} f(n) \quad (f \in \ell^\infty). \quad \square$$

The functional Λ produced above is called a *Banach limit*; the usual notation is $\Lambda(f) := \text{LIM}_{n \rightarrow \infty} f(n)$.

Exercise 10.3. Each Banach limit defines a translation-invariant finitely additive probability measure μ on $\mathcal{P}(\mathbb{N})$ by: $\mu(E) := \text{LIM}_n \chi_E(n)$ for $E \subset \mathbb{N}$.

- (a) Show that $\mu(\{n\}) = 0$ for every $n \in \mathbb{N}$. Conclude that μ is *not* countably additive.
- (b) For n_0 and k in \mathbb{N} , let E denote the arithmetic progression $\{n_0 + kn : n \in \mathbb{N} \cup \{0\}\}$. What is $\mu(E)$?
- (c) Is there an infinite subset E of \mathbb{N} with $\mu(E) = 0$?

This exercise points the “Jekyll and Hyde” character possessed by an infinite dimensional vector space’s algebraic dual. On one hand, the algebraic dual is easy to define and work with (e.g., no worries about continuity). On the other hand, thanks to the Axiom of Choice it has bizarre inhabitants (e.g., Banach limits).

Exercise 10.4 (“Banach limits” for \mathbb{Z} and \mathbb{R}). Show that analogues of “Banach Limit” exist for the additive groups of both the integers and the real line.

10.3 Amenable Groups

Thanks to Corollary 10.4 we know that every abelian group G possesses an *invariant mean*, i.e., a positive linear functional Λ on $B(G)$ that takes value 1 on the constant function 1 and is fixed by the adjoint of every operator of translation by a group element. We’ve noted that such a mean gives rise to a finitely additive probability measure μ on $\mathcal{P}(G)$ that’s *G-invariant* in the sense that $\mu(gE) = \mu(E)$ for each $g \in G$ and $E \in \mathcal{P}(G)$.

Definition 10.11 (Amenable group). To say a group G is *amenable* means that there is a G -invariant, finitely additive probability measure on $\mathcal{P}(G)$, i.e., there is a G -invariant mean on $B(G)$.

Thus every abelian group is amenable. What about the non-abelian ones? Once we venture into the realm of non-commutativity there arises the spectre of “left vs. right.” For non-abelian groups the sort of invariance we’ve been considering should more accurately be called “left-invariance.”

Question. Are there separate notions of “right-” and “left-” amenable?

We’ll see later on (Sect. 12.6) that once a group has a left-invariant mean, then it also has a right-invariant one, and even a “bi-invariant” one. So there are not separate concepts of “left-amenable” and “right-amenable”; it’s all just “amenable.”

Exercise 10.5. Show that every finite group is amenable.

It turns out that *not every* group is amenable. Here's an example, whose apparent simplicity belies its importance:

The Free Group F_2 on Two Generators. The elements of F_2 are “reduced words” of the form $x_1x_2\cdots x_n$ for $n \in \mathbb{N}$ where each x_j comes from the set of symbols $\{a, a^{-1}, b, b^{-1}\}$, subject only to the restriction that no symbol occurs next to its “inverse.” Multiplication in F_2 is defined to be concatenation of words, followed by “reduction,” e.g., $aba^{-1} \cdot abba = abbbba$. Upon allowing the “empty word” e to belong to F_2 we obtain a group.

Caveat: To render the group operation of F_2 “well-defined” it must be shown that the same reduced word results no matter how this reduction is performed. This is not completely trivial (see, e.g., [73, Theorem 1.2, pp. 134–5]). In the next chapter we'll resolve this matter differently by realizing F_2 as a group of rotations of \mathbb{R}^3 .

Exercise 10.6. Convince yourself that (modulo the above *caveat*) F_2 is a group, that it's not abelian, and that it can be visualized as the fundamental group of a figure-eight.

Theorem 10.12. F_2 is not amenable.

Proof. For $x \in \{a, a^{-1}, b, b^{-1}\}$ let $W(x)$ denote the set of reduced words that begin with x . For example, a and $ab^{-1}abb$ belong to $W(a)$, while b and $a^{-1}baab^{-1}$ do not. Thus the sets $W(a), W(a^{-1}), W(b)$, and $W(b^{-1})$ form a pairwise disjoint family of sets in F_2 whose union is $F_2 \setminus \{e\}$. Note that $aW(a^{-1})$ is the set of reduced words in F_2 that don't begin with a , so F_2 is the disjoint union of $W(a)$ and $aW(a^{-1})$; similarly it's also the disjoint union of $W(b)$ and $bW(b^{-1})$.

Now suppose for the sake of contradiction that μ is a finitely additive probability measure on $\mathcal{P}(F_2)$ that is F_2 -invariant. Then, upon using disjointness in the third line below and the invariance of μ in the fourth, we obtain

$$\begin{aligned} 1 &\geq \mu(F_2 \setminus \{e\}) \\ &= \mu(W(a) \cup W(a^{-1}) \cup W(b) \cup W(b^{-1})) \\ &= \mu(W(a)) + \mu(W(a^{-1})) + \mu(W(b)) + \mu(W(b^{-1})) \\ &= \mu(W(a)) + \mu(aW(a^{-1})) + \mu(W(b)) + \mu(bW(b^{-1})) \\ &= \mu(\underbrace{W(a) \cup aW(a^{-1})}_{= F_2}) + \mu(\underbrace{W(b) \cup bW(b^{-1})}_{= F_2}) \\ &= 1 + 1 = 2, \end{aligned}$$

i.e., $1 \geq 2$: a contradiction. □

The question of which groups are amenable is a profound one. We'll see in chapters to come that every solvable group is amenable, but that some compact groups are not. Amenability is intimately connected with the phenomenon of *paradoxicality* which we'll take up in the next chapter; the free group F_2 will play a crucial role.

Notes

The Hahn–Banach Theorem. See, for example, [9, Chap. II, pp. 27–29], [60, Theorem 3.4, p. 21] or [103, Theorem 3.2, pp. 57–58] for the “non-invariant” version. Banach proved a precursor of the Hahn–Banach Theorem in the course of showing that there’s a rotation-invariant mean on $B(\mathbb{T})$ whose resulting finitely additive probability measure on $\mathcal{P}(\mathbb{T})$ is *not* an extension of arc-length measure on the Lebesgue measurable subsets of \mathbb{T} [6, Theorem 19–20]. Banach’s result answered the one dimensional case of a more general problem posed by one of his former professors, Stanisław Ruziewicz.

The Ruziewicz Problem. This problem asks if Lebesgue surface measure on the unit sphere of \mathbb{R}^{n+1} is the unique (up to multiplication by a positive constant) finitely additive, isometry-invariant measure on the Lebesgue measurable subsets of the sphere. The result of Banach mentioned above shows that the answer is “no” for $n = 1$. For $n > 1$ the problem remained open until the 1980s, when the answer was shown to be “yes” by Drinfeld [32] for $n = 2$ and 3, and for $n > 3$ independently by Margulis [74] and Sullivan [114].

The Invariant Hahn–Banach Theorem. This is due to Agnew and Morse [1]; the proof given here is taken from [37, Sects. 3.3 and 3.4].

Banach limits. The result here is due (with a different proof) to Banach [9, Chap. II, p. 34], who also noted the connection with finitely additive probability measures on the subsets of the positive integers [9, Remarques, Sect. 3, p. 231].

Amenable groups. In the 1920s von Neumann [88] initiated the study of groups G for which $\mathcal{P}(G)$ supports invariant finitely additive probability measures. He called such groups “measurable.” The currently preferred term “amenable” was coined by M.M. Day in the late 1950s [28], reputedly as something of a pun on the term “mean” (see [104, p. 34], for example).