Chapter 12
Fixed Points for Non-commuting Map Families

MARKOV–KAKUTANI FOR SOLVABLE FAMILIES

Overview. Here we'll generalize the Markov–Kakutani Theorem (Theorem 9.6, p. 107) to collections of affine, continuous maps that obey a generalized notion of commutativity inspired by the group-theoretic concept of solvability. This will enable us to show, for example, that the unit disc is not paradoxical even with respect to its *full* isometry group, and that solvable groups are amenable, hence not paradoxical. We'll prove that compact solvable groups possess Haar measure, and will show how to extend this result to solvable groups that are just *locally* compact.

12.1 The "Solvable" Markov–Kakutani Theorem

We know from the Banach–Tarski Paradox (Theorem 11.19) that $\mathbb{B}^3$ is paradoxical with respect to the full isometry group of $\mathbb{R}^3$. Thanks to the Markov–Kakutani Fixed-Point Theorem and the commutativity of the group of origin-centered rotations of $\mathbb{R}^2$, we also know (Corollary 10.5, p. 123) that there is defined, for all subsets of the unit disc $\mathbb{B}^2$, a finitely additive probability measure that is rotation-invariant. Consequently (Exercise 11.1) $\mathbb{B}^2$ is not paradoxical with respect to the group of rotations of $\mathbb{R}^2$ about the origin. This raises the question:

Is $\mathbb{B}^2$ paradoxical with respect to its full *group of isometries?*

The isometry group of $\mathbb{B}^2$ allows, in addition to rotations about the origin, *reflections* in a line through the origin; this creates non-commutativity. Indeed, we know from linear algebra that the rotations of $\mathbb{R}^2$ about the origin are the linear transformations represented (with respect to the standard unit-vector basis of $\mathbb{R}^2$) by matrices of the form $\left[\begin{smallmatrix}\cos\theta & -\sin\theta \\ \sin\theta & \cos\theta\end{smallmatrix}\right]$, where $\theta \in [0,2\pi)$ is the angle of rotation. These rotation matrices form the subgroup $\mathrm{SO}(2)$ of $\mathrm{O}(2)$, the group of all 2×2 matrices whose columns form an orthonormal set in $\mathbb{R}^2$. Each matrix in $\mathrm{O}(2)$ has determinant ± 1 (a consequence of column-orthonormality, which can be rephrased: "The transpose of each matrix in $\mathrm{O}(2)$ is its inverse"), those with determinant -1 being the *reflections*

J.H. Shapiro, *A Fixed-Point Farrago*, Universitext,
DOI 10.1007/978-3-319-27978-7_12

about lines through the origin, and those with determinant $+1$ constituting the rotation group $\mathrm{SO}(2)$. The isometries of $\mathbb{R}^2$ that fix the origin are precisely the linear transformations represented by matrices in $\mathrm{O}(2)$. More generally the same is true for $\mathbb{R}^N$, with $\mathrm{O}(N)$ in place of $\mathrm{O}(2)$ (see Appendix D for the full story). Now the matrix group $\mathrm{O}(2)$ (hence its *alter ego*, the isometry group of the unit disc) is not commutative, as witnessed by the pair of matrices:

$$\begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix},$$

the first of which induces rotation through an angle of 45 degrees about the origin, while the second induces reflection about the horizontal axis.

In this chapter we'll generalize the Markov–Kakutani Theorem in a way that applies to non-commutative groups like $\mathrm{O}(2)$; in so doing we'll be able to extend the disc's non-paradoxicality from rotations to *all* its isometries.

Theorem 12.1 (The "Solvable" Markov–Kakutani Theorem)**.** *Suppose K is a non-void compact, convex subset of a Hausdorff topological vector space. Then every* solvable *family of continuous, affine self-maps of K has a common fixed point.*

We'll devote the next section to understanding the meaning of "solvable," after which we'll prove Theorem 12.1 and show how to apply it.

12.2 Solvable Families of Maps

Our notion of solvability is inspired by group theory (see Appendix E).

Definition 12.2. Suppose $\mathscr{A}$ is a family of self-maps of some set.

(a) *Solvable family of maps*. This is what we'll call $\mathscr{A}$ whenever there is a finite chain of subfamilies

$$\{\text{Identity map}\} = \mathscr{A}_0 \subset \mathscr{A}_1 \subset \mathscr{A}_2 \subset \cdots \subset \mathscr{A}_{\mathrm{n}} = \mathscr{A} \tag{12.1}$$

such that for each $1 \le k \le n$ and each pair A, B of maps in $\mathscr{A}_k$ there exists a "commutator" $C \in \mathscr{A}_{k-1}$ such that $AB = BAC$.

(b) *Solvability degree*. More precisely, we may call $\mathscr{A}$ as above "n-solvable."

(c) *Solvable group*. This is what we'll call the family $\mathscr{A}$ whenever it satisfies condition (a) above, and each of the subfamilies $\mathscr{A}_k$ in (12.1) is a *group* under composition. For more precision we may use the term "n-solvable group."

Remarks 12.3. Suppose $\mathscr{A}$ denotes a family of self-maps that is solvable in the sense of Definition 12.2 (a).

(a) *Solvability and commutativity.* $\mathscr{A}_1$ is commutative, so $\mathscr{A}$ is "1-solvable" if and only if it is commutative. "2-solvable" is the next-best thing, … .

(b) *Semigroups and groups of self-maps.* For a family of self-maps of a set, the collection of common fixed points is not changed if one replaces the original family of self-maps by the "unital semigroup" it generates, i.e., the set of all possible finite compositions of the original maps, along with the identity map. If each map of the original family is a bijection, we can even add all the inverses to the original family without changing the common fixed-point set, in which case the new "inverse-enhanced" family generates a *group* under composition having the same common fixed-point set as the original family. Here we'll only consider self-map families that are groups.

(c) *Solvable groups.* Suppose G is a group with identity element e. We can consider G to be a group of self-maps, acting itself by (say) left multiplication. Each pair of elements $a, b \in G$ has a unique commutator $[a;b] := (ba)^{-1}ab = a^{-1}b^{-1}ab$. Thus, according to Definition 12.2(c) above: G is a *solvable group* if and only if there is a chain of *subgroups*

$$\{e\} = G_0 \subset G_1 \subset G_2 \ldots G_n = G \tag{12.2}$$

such that for each index k between 1 and n the subgroup G_{k-1} contains all the commutators of G_k.

The usual definition of "solvable" for groups stipulates that for $1 \le k \le n$ the subgroup G_{k-1} must to be a *normal* subgroup of G_k, and that furthermore each quotient group G_k/G_{k-1} must be abelian. These requirements of normality plus commutativity turn out to be equivalent to the single commutator-containment condition of the last paragraph; see Appendix E for the details.

Example 12.4. *The matrix group* $\mathrm{O}(2)$ *is solvable.* We've noted that the family of isometric self-maps of $\mathbb{B}^2$ can be identified with $\mathrm{O}(2)$ acting by left-multiplication of column vectors. Consider the chain of subgroups

$$\{I\} \subset \mathrm{SO}(2) \subset \mathrm{O}(2), \tag{12.3}$$

noting that $\mathrm{SO}(2)$, the group of 2×2 rotation matrices, is commutative. The multiplicative property of determinants now takes over; each matrix in $\mathrm{O}(2)$ has determinant either $+1$ or -1, and so has the same determinant as its inverse. Thus given matrices A and B in $\mathrm{O}(2)$ the commutator $[A;B]$ belongs to $\mathrm{O}(2)$ and has determinant $+1$; it therefore belongs to $\mathrm{SO}(2)$.

Conclusion: $\mathrm{O}(2)$ is a 2-solvable group in the sense of Definition 12.2.

Example 12.5. *The affine group of* $\mathbb{R}$ *is solvable.* Let $A(\mathbb{R})$ denote the collection of affine transformations of the real line, i.e., the transformations $\gamma_{r,t} : x \to rx + t$ $(x \in \mathbb{R})$ for $t \in \mathbb{R}$ and $r \in \mathbb{R}\backslash\{0\}$. Then, with composition as the binary operation in $A(\mathbb{R})$:

(a) $\gamma_{r,t} \circ \gamma_{\rho,\tau} = \gamma_{r\rho, r\tau+t}$, so $A(\mathbb{R})$ is a group, with $\gamma_{r,t}^{-1} = \gamma_{1/r,-t/r}$.

(b) $A(\mathbb{R})$ is generated by two commutative subgroups: the dilation group consisting of maps $\gamma_{r,0}$ where $r \neq 0$, and the translation group $T(\mathbb{R})$ consisting of maps $\gamma_{1,t}$ for $t \in \mathbb{R}$.

(c) The commutator $[\gamma_{r,t};\gamma_{\rho,\tau}] = \gamma_{1,s} \in T(\mathbb{R})$, where $s = \frac{(1-\rho)t-(1-r)\tau}{r\rho}$.

Thus we have the chain of groups: $\{\text{identity map}\} \subset T(\mathbb{R}) \subset A(\mathbb{R})$, where $T(\mathbb{R})$ is commutative and contains all the commutators of $A(\mathbb{R})$.

Conclusion: $A(\mathbb{R})$ is a 2-solvable group.

Exercise 12.1. Show that the map $\gamma_{r,t} \to \left[\begin{smallmatrix} r & t \\ 0 & 1 \end{smallmatrix}\right]$ is a homomorphism taking $A(\mathbb{R})$ onto a group of invertible 2×2 real matrices, and that all the calculations in the example above can be done "matricially."

The exercises below give two more examples of solvable matrix groups, the second of which is 3-solvable, but not 2-solvable.

Exercise 12.2 (The Heisenberg group is solvable). The *Heisenberg group* is the collection $\mathscr{H} = \mathscr{H}_3(\mathbb{R})$ of 3×3 real matrices that are upper triangular and whose main diagonal consists entirely of 1's.

(a) Show that $\mathscr{H}$ is a group under matrix multiplication.

(b) Let $\mathscr{K}$ denote the subset of $\mathscr{H}$ consisting of matrices of the form $\left(\begin{smallmatrix} 1 & 0 & a \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{smallmatrix}\right)$. Show that $\mathscr{K}$ is a commutative subgroup of $\mathscr{H}$.

(c) Show that if $A, B \in \mathscr{H}$, then the commutator $A^{-1}B^{-1}AB$ belongs to $\mathscr{K}$. Conclude that $\mathscr{H}$ is 2-solvable.

Exercise 12.3 (The Upper-Triangular group is 3-solvable, but not 2-solvable). Let $\mathscr{U}$ denote the collection of 3×3 matrices that are *upper triangular,* i.e., have all entries zero below the main diagonal.

(a) Show that $\mathscr{U}$ is a group under matrix multiplication.

(b) Show that the Heisenberg group contains every $\mathscr{U}$-commutator. Conclude that $\mathscr{U}$ is 3-solvable.

(c) Show that the $\mathscr{U}$ is not 2-solvable.

(d) *Suggestion:* By considering, e.g., matrices of the form $A = \left(\begin{smallmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & a \end{smallmatrix}\right)$ and $B = \left(\begin{smallmatrix} b & 0 & c \\ 0 & 1 & 0 \\ 0 & 0 & d \end{smallmatrix}\right)$, show that the collection of commutators of $\mathscr{U}$ exhausts the *entire* Heisenberg group. Argue that if $\mathscr{U}$ were 2-solvable, then the Heisenberg group would have to be commutative, which it is not.

The next exercise concerns a famous class of finite groups that are *not* solvable.

Exercise 12.4. *S_n is not solvable for $n \geq 5$.* Here S_n denotes the set of permutations (1-to-1 onto maps) of a set of n elements, which we might as well take to be $[1,n] := \{1,2,\dots n\}$. With composition as its binary operation, S_n is a group (the symbol "S" stands for "symmetric"). We assume here that $n \geq 5$.

Of particular interest to us are the *3-cycles* in S_n, i.e., the maps that permute a triple $\{a,b,c\}$ of distinct elements of $[1,n]$ *cyclically:* $a \to b \to c \to a$ and leave everything else alone. Notation for such a 3-cycle: (a,b,c). In the exercises below we assume $n \geq 5$.

(a) Show that the 3-cycle $(1,4,3)$ is the commutator $[\sigma,\tau]$ where $\sigma = (1,2,3)$ and $\tau = (3,4,5)$.

(b) By making appropriate substitutions in part (a) show that *every* 3-cycle in S_n is a commutator of other 3-cycles.

(c) Use part (b) to show that S_n is not solvable.

12.3 Proof of the solvable Markov–Kakutani Theorem

The argument proceeds by induction on the "solvability index" n in (12.1). Since $\mathscr{A}_1$ is commutative, the case $n = 1$ is just the original Markov–Kakutani Theorem (Theorem 9.6, p. 107).

For the induction step suppose $n \geq 2$ and the result is true for all $(n-1)$-solvable families, of which $\mathscr{A}_{n-1}$ in (12.1) is one. The set K_{n-1} of common fixed points for $\mathscr{A}_{n-1}$ is nonempty (induction hypothesis), compact (continuity of the maps in $\mathscr{A}_{n-1}$), and convex (affine-ness of the maps in $\mathscr{A}_{n-1}$).

Claim: Each map $A \in \mathscr{A} = \mathscr{A}_n$ takes K_{n-1} into itself.

Proof of Claim. Given $A \in \mathscr{A}$ and $p \in K_{n-1}$ we're claiming that $A(p) \in K_{n-1}$, i.e., that $BA(p) = A(p)$ for every $B \in \mathscr{A}_{n-1}$. Given $A \in \mathscr{A}$ and $B \in \mathscr{A}_{n-1}$ there exists $C \in \mathscr{A}_{n-1}$ such that $BA = ABC$. Thus for $p \in K_{n-1}$ we have (since both B and C belong to $\mathscr{A}_{n-1}$): $BA(p) = ABC(p) = AB(p) = A(p)$, as desired.

To finish the proof of Theorem 12.1 we're going to show that $\tilde{\mathscr{A}}$, the collection of restrictions to K_{n-1} of maps in $\mathscr{A}$, is commutative. This, along with the just-proved *Claim*, will establish $\tilde{\mathscr{A}}$ as a commutative family of continuous, affine self-maps of K_{n-1}. The original Markov–Kakutani Theorem will then provide for $\tilde{\mathscr{A}}$ a common fixed point $p \in K_{n-1}$, *a fortiori* a fixed point for every map in $\mathscr{A}$.

It remains to establish the desired commutativity for $\tilde{\mathscr{A}}$. For this, suppose A and B belong to $\mathscr{A}$ and choose $C \in \mathscr{A}_{n-1}$ so that $AB = BAC$. Then for $p \in K_{n-1}$ (hence a fixed point for C): $A(B(p)) = B(A(C(p))) = B(A(p))$, i.e., $AB = BA$ on K_{n-1}. □

12.4 Applying the solvable M–K Theorem

Recall the cast of characters that emerged in Chaps. 9 and 10 when we applied the original the Markov–Kakutani Theorem.

(a) There was a set S and a commutative family Φ of self-maps of S.

(b) Each $\varphi \in \Phi$ gave rise to the (linear) composition operator $C_\varphi : f \to f \circ \varphi$ acting on $B(S)$ (the vector space of bounded, real-valued functions on S). We denoted the collection of all such composition operators by C_Φ.

These actors will return in this chapter, except that now we'll allow Φ to be "solvable" in the sense of Definition 12.2. The Markov–Kakutani triple $(X, K, \mathscr{A})$ of Sects. 9.5 and 10.1 will return unchanged:

(c) $X = B(S)^\sharp$, the algebraic dual of $B(S)$, taken in its *weak-star topology*.

(d) $K = \mathscr{M}(S)$, the set of "means" on $B(S)$, i.e., those positive linear functionals on $B(S)$ that take value 1 on the function $\equiv 1$ on S.
(e) $\mathscr{A} = C_\Phi^\sharp$, the collection of adjoints of composition operators belonging to the family C_Φ.

To apply our enhanced Markov–Kakutani Theorem we need only to show that the solvability assumed for the original family Φ of self-maps of S is inherited by the family $C_\Phi^\sharp$ of affine self-maps of $\mathscr{M}(S)$. For this one need only check that the map $\varphi \to C_\varphi$ reverses composition ($C_{\varphi\circ\psi} = C_\psi C_\varphi$), and that the same is true of the map $T \to T^\sharp$ that associates to each linear transformation on a vector space its adjoint. Thus the map $\varphi \to C_\varphi^\sharp$ preserves the order of composition; in particular, if $\varphi, \psi, \gamma \in \Phi$ and γ is a commutator of the pair (φ, ψ) in the sense that $\varphi \circ \psi = \psi \circ \varphi \circ \gamma$, then $C_\gamma^\sharp$ is a commutator of the pair $(C_\varphi^\sharp, C_\psi^\sharp)$. Consequently, if the original family of maps Φ is solvable, then so is $C_\Phi^\sharp$. Theorem 12.1 can therefore be applied to yield

Theorem 12.6 (Invariant means for solvable families of maps)**.** *Suppose Φ is a solvable family of self-maps of a set S. Then:*

(a) *There is a mean Λ on $B(S)$ that is invariant for $C_\Phi^\sharp$, i.e., $\Lambda \circ C_\varphi = \Lambda$ for every $\varphi \in \Phi$ (cf. Theorem 9.19).*
(b) *There is a finitely additive Φ-invariant probability measure on $\mathscr{P}(S)$ (cf. Theorem 10.3).*
(c) *S is not Φ-paradoxical (see Exercise 11.1).*

Since $\Phi = \mathrm{O}(2)$ is a solvable family of self-maps of $\mathbb{B}^2$ and S^1, we see in particular:

Corollary 12.7. *$\mathbb{B}^2$ and S^1 are not $\mathrm{O}(2)$-paradoxical.*

Corollary 12.8. *Solvable groups are amenable, hence not paradoxical.*

In the other direction we have

Corollary 12.9. *The following groups are not solvable:*

(a) *The free group F_2 on two letters.*
(b) *The compact group $\mathrm{SO}(3)$ of 3×3 orthogonal matrices with determinant one.*

Proof. Both groups are not amenable (Theorem 10.12 for F_2 and Corollary 11.9 for $\mathrm{SO}(3)$ in its guise as the rotation group $\mathscr{R}$), hence not solvable. □

If our basic set S is a compact topological space, then we have the following extension of Corollaries 9.20 and 9.21:

Corollary 12.10. *If Φ is a solvable family of continuous affine self-maps of a compact topological space S, then there exists a regular Borel probability measure μ for S such that*

$$\int f \circ \varphi \, d\mu = \int f \, d\mu$$

for every $f \in C(S)$ and $\varphi \in \Phi$.

Corollary 12.11. *Every solvable compact topological group has a Haar measure.*

12.5 The (solvable) Invariant Hahn–Banach Theorem

Given a solvable family Φ of *continuous* self-maps of a *compact Hausdorff space S*, our solvably enhanced version of the Markov–Kakutani Theorem produces—just as did the original version in Sect. 10.2—two important Φ-invariant set functions for S: A regular probability measure μ on the Borel sets of S (Corollary 12.10), and a finitely additive probability measure ν defined for all subsets of S (Theorem 12.6). This brings up the same question we faced in Sect. 10.2: "Can ν be realized as an extension of μ?" Once again the answer is "yes," with the heavy lifting done by a "solvable" extension of the Invariant Hahn–Banach Theorem (Theorem 10.6, p. 124).

Theorem 12.12 (The "solvable" Invariant Hahn–Banach Theorem)**.** *Suppose V is a vector space and $\mathscr{G}$ is a solvable family of linear transformations $V \to V$. Suppose W is a linear subspace of V that is taken into itself by every transformation in $\mathscr{G}$, and that p is a gauge function on V that is "$\mathscr{G}$-subinvariant" in the sense that*

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

Suppose Λ is a $\mathscr{G}$-invariant functional on W that is dominated by p, i.e.,

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

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

Corollary 12.13. *If S is a compact Hausdorff space upon which acts a solvable family Φ of continuous self-maps, then each regular Φ-invariant probability measure on the Borel sets of S extends to a finitely additive probability measure defined for* all *subsets of S.*

The proofs of these two results are identical to those of their commutative analogues (Theorem 10.6 and Corollary 10.7, pp. 124–125), except that the solvable Markov–Kakutani Theorem replaces the original one.

We saw in Corollary 10.5 that for the closed unit disc $\mathbb{B}^2$ there is a finitely additive, rotation-invariant probability measure on $\mathscr{P}(\mathbb{B}^2)$. Thanks to Example 12.4 and Theorem 12.1 we now know there exists a finitely additive probability measure on $\mathscr{P}(\mathbb{B}^2)$ invariant for the *full* isometry group $\mathrm{O}(2)$ of $\mathbb{B}^2$. Corollary 12.13 shows that this isometry-invariant finitely additive probability measure can be chosen to extend normalized Lebesgue area measure; similar results hold for the unit circle.

Invariant extension of Lebesgue measure on $\mathbb{R}$. Lebesgue measure m on the Borel subsets of the real line is invariant under translations, and "scales properly" under dilations. More precisely: for each pair (r,t) of real numbers, and each Borel subset E of $\mathbb{R}$, we have $m(rE+t) = |r|m(E)$. Thanks to the solvability of the affine group $A(\mathbb{R})$ of the real line (Exercise 12.5), our "solvable" Invariant Hahn–Banach Theorem provides an extension of Lebesgue measure to a finitely additive measure on $\mathscr{P}(\mathbb{R})$ that preserves the translation-invariance and scaling properties of the original. More precisely:

Theorem 12.14. *There is an extension of Lebesgue measure to a finitely additive measure μ on all subsets of $\mathbb{R}$ such that for each $E \subset \mathbb{R}$:*

$$\mu(rE+t) = |r|\mu(E) \qquad (0 \neq r \in \mathbb{R},\ t \in \mathbb{R}) \tag{12.4}$$

and

$$m_*(E) \leq \mu(E) \leq m^*(E), \tag{12.5}$$

where $m_*(E)$ and $m^*(E)$ denote, respectively, the inner and outer measures of E.

Proof. Let V be the collection of real-valued functions f on $\mathbb{R}$ for which the upper integral

$$\int^* |f| = \inf\left\{\int s : s \in \mathscr{S}, |f| \leq s\right\}$$

of $|f|$ over $\mathbb{R}$ is finite; here $\mathscr{S}$ denotes the collection of Borel-measurable, integrable functions on $\mathbb{R}$ that are *simple*, i.e., take only finitely many values, and the integrals are taken with respect to Lebesgue measure on the line. For $f \in V$ let $p(f) = \int^* |f|$.

Exercise 12.5. Prove that p is a gauge function on V, as defined in the statement of the Hahn–Banach Theorem on p. 124.

For $\gamma = \gamma_{r,t} \in A(\mathbb{R})$ (notation as in Example 12.5), define the linear transformation L_γ on V as the "weighted" composition operator:

$$(L_\gamma f)(x) = rf(\gamma(x)) \qquad (f \in V,\ x \in \mathbb{R}),$$

and let $\mathscr{A}$ denote the collection of all such transformations. With composition as its binary operation, $\mathscr{A}$ is a group that inherits the 2-solvability of $A(\mathbb{R})$, and so satisfies the hypotheses of the "solvable" Invariant Hahn–Banach Theorem.

Thanks to the change-of-variable formula for Lebesgue integrals, the functional p is invariant for each $L_\gamma \in \mathscr{A}$:

$$p(L_\gamma f) = p(f) \qquad (f \in V,\ \gamma \in A(\mathbb{R})).$$

Let W denote the subspace of V consisting of functions whose absolute value is Lebesgue measurable, and so Lebesgue integrable. On W let λ be the linear functional of integration with respect to Lebesgue measure m. Then λ, too, is $\mathscr{A}$-invariant so Theorem 12.12 provides a $\mathscr{A}$-invariant linear functional Λ on V that extends λ and is dominated on V by p.

Now for the desired finitely additive measure: if E is a subset of $\mathbb{R}$ with finite outer measure then its characteristic function χ_E is in V (its upper integral is precisely $m^*(E)$), so we can set $\mu(E) = \Lambda(\chi_E)$. The $\mathscr{A}$-invariance of Λ translates into property (12.4) for μ, while the fact that $\Lambda \leq p$ on V shows us that

$$\int_* f = -p(-f) \leq \Lambda(f) \leq p(f) = \int^* f \qquad (f \in V),$$

where on the right we see the *lower integral* of f, i.e., the supremum of the integrals of integrable simple functions that are $\leq f$ at each point of $\mathbb{R}$. In particular, for $f = \chi_E$ with $m^*(E) < \infty$ we obtain (12.5).

It remains only to extend μ to *all* subsets of $\mathbb{R}$, which we do by defining $\mu(E) = \infty$ whenever $m^*(E) = \infty$. With the usual conventions involving arithmetic with ∞, the result is still a finitely additive measure that preserves the desired properties. □

Higher dimensional extensions? Does the above result extend to affine maps of $\mathbb{R}^N$ for $N > 1$? In this case the maps are $\gamma_{A,v} : x \to Ax + v$ with A in the group of invertible $n \times n$ real matrices and v a vector in $\mathbb{R}^N$. The change of variable formula now tells us that $\lambda(\gamma_{A,v}(E)) = \det(A)\lambda(E)$ for each Borel subset E of $\mathbb{R}^N$, where λ denotes Lebesgue measure on $\mathbb{R}^N$, so the question is: For $n > 1$ does there exist a finitely additive extension of Lebesgue measure to all the subsets of $\mathbb{R}^N$ that satisfies the above transformation formula.

The answer is "No!" For $N = 3$ the Banach–Tarski Paradox tells us that no such measure exists, even for the subgroup of $A(\mathbb{R}^3)$ consisting of isometries of $\mathbb{R}^3$. The Banach–Tarski Paradox extends to $\mathbb{R}^N$ with $N > 3$ (the proof is an adaptation—not entirely trivial—of the three dimensional one; see, for example, [121, Chap. 5]), with the same result for extensions of Lebesgue measure. For $N = 2$ there is no Banach–Tarski Paradox to help us out here. In its place, however, is the *von Neumann Paradox*, according to which any two bounded subsets of $\mathbb{R}^2$ with nonvoid interior are equidecomposable with respect to the group of affine maps $\gamma_{A,v}$ for which $\det(A) = 1$, i.e., the group of *area-preserving* affine maps. Thus, once again there is no hope for a two dimensional extension of Theorem 12.14.

Countably additive extensions? The usual construction of a subset of $\mathbb{R}$ that's not Lebesgue measurable shows that (assuming the Axiom of Choice) there is no countably additive extension of Lebesgue measure to all subsets of the real line.

12.6 Right vs. Left

Having left the friendly confines of commutativity, we need to address the question of "rightness vs. leftness" for invariant Borel measures on topological groups, and more generally for means on "non-topological" groups (recall Definition 9.16, p. 114). To this point "invariant," for a group G and a mean Λ on $B(G)$ has meant that $L_\gamma^\sharp \Lambda := \Lambda \circ L_\gamma = \Lambda$ for each of the "left-translation maps" $L_\gamma : B(G) \to B(G)$ defined for $\gamma \in G$ by

$$L_\gamma(f)(x) = f(\gamma x) \qquad (x \in G, f \in B(G)). \tag{12.6}$$

If a *compact* group G has such an invariant mean (e.g., if G is abelian, or more generally, solvable) then the Riesz Representation Theorem (Sect. 9.2, p. 104) associates with the restriction of this mean to $C(G)$ a similarly invariant regular Borel probability measure—a Haar measure.

For non-commutative groups we need to address the corresponding idea of "right-invariance" that utilizes the transformations $R_\gamma\colon B(G) \to B(G)$ defined by

$$R_\gamma(f)(x) = f(x\gamma) \qquad (x \in G, f \in B(G)). \tag{12.7}$$

The question arises: "Is right-invariance the same as left invariance?"

Right vs. left Haar measure. We'll see in the next chapter that every compact topological group has a left-invariant regular Borel probability measure; we know right now (Corollary 12.11 above) that such a measure exists if the group is *solvable.* For the time being, however, let's just assume the existence of such a measure for a given compact group and see where this leads.

Theorem 12.15. *Suppose G is a compact topological group and μ is a left-invariant regular Borel probability measure for G. Then:*

(a) *μ is also right-invariant, hence "bi-invariant."*
(b) *There is no other invariant regular probability measure for G.*
(c) *μ is "inversion invariant":*

$$\int f(x^{-1})\,d\mu(x) = \int f\,d\mu \qquad (f \in C(G)),$$

i.e., $\mu(B) = \mu(B^{-1})$ for every Borel subset B of G.

Proof. Suppose ν is a right-invariant regular probability measure for G. Then for each $f \in C(G)$ the left-invariance of μ demands that

$$\int f\,d\mu = \int f(xy)\,d\mu(y) \qquad (x \in G),$$

hence

$$\begin{aligned}
\int f\,d\mu &= \int \left(\int f(xy)\,d\mu(y)\right) d\nu(x) && [\nu(G) = 1] \\
&= \int \left(\int f(xy)\,d\nu(x)\right) d\mu(y) && [\text{Fubini}] \\
&= \int \left(\int f(x)\,d\nu(x)\right) d\mu(y) && [\nu \text{ right invariant}] \\
&= \int f\,d\nu && [\mu(G) = 1].
\end{aligned}$$

Thus $\mu = \nu$, which establishes (a) and (b).

As for (c), note that μ has a natural right-invariant companion $\tilde{\mu}$ defined, thanks to the Riesz Representation Theorem, by

$$\int f\,d\tilde{\mu} = \int f(x^{-1})\,d\mu(x) \qquad (f \in C(G)).$$

By (b) we must have $\tilde{\mu} = \mu$, thus establishing the inversion-invariance of μ. □

Summary: For *compact* topological groups there's no distinction between left- and right-invariant Borel probability measures. Such a "bi-invariant" measure (whose existence we'll prove in the next chapter) is unique, and even "inversion-invariant."

In the next section we'll discuss topological groups that are *not* compact. The following exercise shows that in this generality there may be left-invariant Borel measures that are not right-invariant.

Exercise 12.6 (Haar measure(s) on a non-compact group). Let $G = \{(x,y) \in \mathbb{R}^2 : x > 0\}$ be the open right-half plane of $\mathbb{R}^2$ with the binary operation:

$$(a,b)\cdot(x,y) := (ax, ay+b) \qquad ((a,b),(x,y) \in G).$$

(a) Show that G, in the operation described above, is a topological group that is *solvable*. *Suggestion:* Examine the map $(x,y) \to \left[\begin{smallmatrix} x & y \\ 0 & 1 \end{smallmatrix}\right]$ (cf. Example 12.5 and Exercise 12.1).

(b) Use the change-of-variable formula for double integrals to show that the measure $dx\,dy/x^2$ is left-invariant on G, but not right-invariant.

(c) Show that the measure $dx\,dy/x$ is right-invariant on G, but not left-invariant.

Right- vs. left-invariant means. For non-abelian groups the situation of left- vs. right-invariance of *means* is more subtle than the one described above for measures. It turns out that left-invariant means need not be right-invariant (and vice versa), but once there is a left- or right-invariant mean, there is a "bi-invariant one." Thus there is no "left vs. right" problem with the notion of "amenable."

In addition to the notions of left and right invariance for means, there is a notion of *inversion-invariance* that mirrors the property observed for invariant measures in Theorem 12.15. Define the linear transformation $J\colon B(G) \to B(G)$ by

$$(Jf)(x) = f(x^{-1}) \qquad (f \in B(G),\ x \in G) \tag{12.8}$$

and call a mean Λ on $B(G)$ *inversion invariant* if $J^\sharp\Lambda = \Lambda$, i.e., if $\Lambda \circ J = \Lambda$.

Theorem 12.16. *Suppose G is a group for which $B(G)$ has a left-invariant mean. Then $B(G)$ has a mean that is both bi-invariant and inversion invariant.*

Proof. We'll first show that every left-invariant mean has a right-invariant counterpart. To this end note that for the inversion operator J defined by (12.8),

$$R_\gamma J = JL_{\gamma^{-1}} \qquad \text{and} \qquad L_\gamma J = JR_{\gamma^{-1}}\,. \tag{12.9}$$

Thus if λ is a left-invariant mean for $B(G)$ then the calculation below shows that $\rho = J^\sharp\lambda$ is right invariant. For every $\gamma \in G$:

$$R_\gamma^\sharp \rho \;=\; R_\gamma^\sharp J^\sharp \lambda \;=\; (JR_\gamma)^\sharp\lambda \;=\; (L_{\gamma^{-1}}J)^\sharp\lambda \;=\; J^\sharp L_{\gamma^{-1}}^\sharp \lambda \;=\; J^\sharp\lambda \;=\; \rho,$$

where the middle equality above comes from the first identity of (12.9) and the next-to-last one from the left-invariance of λ. Clearly $\rho(1) = 1$, and it's easy to check that ρ is a positive linear functional on $B(G)$, hence a right-invariant mean.

From a left-invariant mean λ on $B(G)$ and a right-invariant one ρ, the definition below provides a *bi-invariant* one: a mean ν on $B(G)$ with $L_\gamma^\sharp \nu = R_\gamma^\sharp \nu = \nu$ for every $\gamma \in G$:

$$\nu(f) = \lambda(\tilde{f}) \quad \text{where} \quad \tilde{f}(\gamma) = \rho(L_\gamma f) \qquad (f \in B(G), \gamma \in G).$$

One checks easily that ν is a mean on $B(G)$. As for its bi-invariance, a little calculation (using the identity $L_{\gamma\beta} = L_\beta L_\gamma$, and the fact that every left translation commutes with every right translation) shows that for each $\gamma \in G$ and $f \in B(G)$:

$$\widetilde{R_\gamma f} = \tilde{f} \quad \text{and} \quad \widetilde{L_\gamma f} = L_\gamma \tilde{f},$$

whereupon

$$\nu(R_\gamma f) = \lambda(\widetilde{R_\gamma f}) = \lambda(\tilde{f}) = \nu(f)$$

and

$$\nu(L_\gamma f) = \lambda(\widetilde{L_\gamma f}) = \lambda(L_\gamma \tilde{f}) = \lambda(\tilde{f}) = \nu(f)$$

as desired.

Finally, from the bi-invariant mean ν we form an inversion-invariant one $\eta = (\nu + J^\sharp \nu)/2$ that is easily to inherit the bi-invariance of ν. □

Example 12.17 (A left-invariant mean that's not right-invariant). Let G be the group of Exercise 12.6; the identity of this group is the point $(1,0)$, and the inverse of $(x,y) \in G$ is $(1/x, -y/x)$.

By part (a) of Exercise 12.6 we know that G is solvable, hence Corollary 12.12, our "solvable" Invariant Hahn–Banach Theorem, applies to $B(G)^\sharp$. In particular, let p denote the gauge function on $B(G)$ defined by the iterated upper limits

$$p(f) = \limsup_{y \to \infty} \left[\limsup_{x \to \infty} f(x,y) \right] \qquad (f \in B(G)).$$

Let W denote the set of all functions $f \in B(G)$ for which the iterated limit

$$\lambda(f) = \lim_{y \to \infty} \left[\lim_{x \to \infty} f(x,y) \right]$$

exists (finitely). One checks easily that:

(a) p is left-invariant on $B(G)$: $p \circ L_\gamma = p$ for every $\gamma \in G$.
(b) W is a linear subspace of $B(G)$ with $L_\gamma(W) \subset W$ for each $\gamma \in G$, and
(c) λ is a linear functional on W that is left-invariant for G.

Since G is solvable, our extended Invariant Hahn–Banach Theorem applies to produce an extension of λ to a left-invariant linear functional Λ on $B(G)$.

However Λ is *not* right-invariant for G. For example, if $g(x,y) := xy/(x^2 + y^2)$ then $g \in W$ with $\lambda(g) = 0$. For $(a,b) \in G$ we have for each $y \in \mathbb{R}$:

$$\lim_{x\to\infty} R_{(a,b)}g(x,y) = \lim_{x\to\infty} \frac{ax(bx+y)}{(ax)^2+(bx+y)^2} = \frac{ab}{a^2+b^2} = g(a,b).$$

Thus $\lambda(R_{(a,b)}g) = g(a,b)$ which is $\neq 0$ if $b \neq 0$. However $\lambda(g) = 0$, so the functional λ, and therefore its extension Λ, is not right-invariant.

Exercise 12.7. Show that the mean Λ of Example 12.17 is not inversion-invariant.

Exercise 12.8 (Banach limits for solvable groups). Suppose G is an infinite group that is *solvable*. For a function $f : G \to \mathbb{R}$ and $c \in \mathbb{R}$ define "$\lim_{\gamma\to\infty} f(\gamma) = c$" to mean: "For every $\varepsilon > 0$ there exists a finite subset F_ε of G such that $|f(\gamma) - c| < \varepsilon$ for every $\gamma \in G\backslash F_\varepsilon$." Make similar definitions for upper and lower limits. Use Corollary 12.12 to show that there exists a mean Λ on $B(G)$ that is both bi-invariant and inversion-invariant for G, and for which

$$\liminf_{\gamma\to\infty} f(\gamma) \leq \Lambda(f) \leq \limsup_{\gamma\to\infty} f(\gamma)$$

for each $f \in B(G)$.

12.7 The Locally Compact Case

To say a topological space X is *locally compact* means that for each point $x \in X$, every neighborhood of x contains a compact neighborhood of x. In other words, at each point the topology of the space has a local base of compact neighborhoods. As we've mentioned previously (but will not prove here), every locally compact group has a left—and therefore a right—Haar measure. Under appropriate regularity conditions left Haar measure is unique up to positive scalar multiples, as is right Haar measure, but we've already seen (Exercise 12.6) that left and right Haar measures need not be scalar multiples of each other. Detailed proofs of the existence and uniqueness of Haar measure on locally compact groups exist in many places; see, for example, [39, Chap. 2] or [29, Chap. 7]. There does not, however, seem to be a neat functional-analysis proof of this result. The purpose of this section is to show how our "Markov–Kakutani method" can be modified to provide Haar measure, at least for locally compact groups that are *solvable*.

Throughout this discussion it will help to keep in mind three examples: Lebesgue measure on Euclidean space, and the left and right Haar measures on the group G of Exercise 12.6. All three measures are unbounded, and the last two show that, even in the solvable case, left and right Haar measures can be essentially different.

Regular and Radon measures. Suppose μ is a Borel measure for a locally compact (Hausdorff) space X. To say that a Borel set $E \subset X$ is:

- μ*-outer regular* means that $\mu(E) = \inf\{\mu(U) : U \text{ is open and } U \supset E\}$.
- μ*-inner regular* means $\mu(E) = \sup\{\mu(K) : K \text{ is compact and } K \subset E\}$.

To say that μ itself is

- *Regular* means that every Borel set is both inner and outer regular (similarly we can attach to μ the terms "inner regular" or "outer regular").

- *Locally finite* means that every point of the space has a neighborhood of finite measure.
- *A Radon measure* means that it is locally finite, outer regular, and every *open set* is μ-inner regular.

We'll see below (Exercise 12.10) that in general not every Radon measure is regular. This is not something to worry about, especially in the compact case:

Exercise 12.9. Show that for *compact* spaces every Radon measure is regular.

Haar Measure To say that μ is a *left Haar measure* for a locally compact group G means that μ is a Radon measure for G that is invariant for left-translation on the group, i.e., $\mu(E) = \mu(gE)$ for every Borel subset E of G and every $g \in G$. Right Haar measure is defined similarly.

We'll need the full-strength version of the Riesz Representation Theorem. If X is a locally compact space that is not compact, the space $C(X)$ of all continuous real-valued functions on X is no longer an appropriate setting for the Riesz theorem; non-compact spaces raise the spectre of unbounded functions and infinite measures, creating problems for the integration of arbitrary continuous functions against arbitrary Borel measures. The resolution is to replace $C(X)$ by its subspace $C_c(X)$: those continuous functions on X that have *compact support*, i.e., that vanish off some compact set. Each such function is bounded and can be integrated against every locally finite Borel measure. Such measures therefore induce linear functionals on $C_c(X)$, the positive ones inducing positive functionals. The Riesz Representation Theorem says that each positive linear functional on $C_c(X)$ is given by integration against such a measure and, with the appropriate conditions of regularity, this representing measure is unique. More precisely (see, e.g., [102, Theorem 2.14, pp. 40–41]):

The Riesz Representation Theorem for locally compact spaces. *Suppose X is a locally compact space and Λ a positive linear functional on $C_c(X)$. Then there is a unique Radon measure for X such that $\Lambda(f) = \int f\,d\mu$ for every $f \in C_c(X)$.*

The exercise below shows that Haar measure—even for a commutative locally compact group—need not always be regular. The group G in question is the additive group $\mathbb{R}^2$ endowed with the product topology it gets when viewed as $\mathbb{R}_d \times \mathbb{R}$, where $\mathbb{R}_d$ denotes the real line with the *discrete* topology.

Exercise 12.10 (A non-regular Haar measure). For the group G described above:

(a) Show that G is locally compact, and even *metrizable* (for $p_j = (x_j, y_j) \in G (j = 1,2)$ take $d(p_1, p_2)$ equal to $|y_1 - y_2|$ if $x_1 = x_2$, and $1 + |y_1 - y_2|$ otherwise).

(b) Let δ denote the counting measure for $\mathbb{R}_d$ and let λ denote Lebesgue measure on the Borel subsets of $\mathbb{R}$. Show that each of these is a Haar measure for its respective topological group.

(c) Show that $\mathbb{R}_d \times \{0\}$ (the "discrete x-axis" has μ-measure ∞, whereas each of its compact subsets has μ-measure zero. Thus μ is not regular. Show that μ is, nevertheless, a Radon measure.

(d) Show that μ is a Haar measure for G.

Here is the main result of this section.

Theorem 12.18. *Every locally compact solvable topological group has a Haar measure.*

Strategy of proof. Assume G is solvable and locally compact, but not compact. Let e denote the identity element of G. We seek to produce left Haar measure for G by applying our "solvable" Markov–Kakutani theorem to an appropriate subset of the algebraic dual $C_c(G)^\sharp$ of $C_c(G)$. Since non-zero constant functions no longer belong to $C_c(G)$, the space of "means" that worked so well in the compact situation no longer exists. Our argument will hinge on finding a substitute.

Simplifying assumptions. To keep the argument as transparent as possible we'll assume G is *metrizable* and that its metric d is *G-invariant* in the sense that $d(\gamma x, \gamma y) = d(x,y)$ for all $x, y, \gamma \in G$ (i.e., for each $\gamma \in G$ the left-translation map $L_\gamma : x \to \gamma x$ $(x \in G)$ is an isometry). See the *Notes* at the end of this chapter for some discussion of these assumptions.

Notation. Let $B_r(x)$ denote the open d-ball of radius $r > 0$, centered at $x \in G$. The G-invariance of d insures that $\gamma B_r(x) = B_r(\gamma x)$ for all $x, \gamma \in G$. For $f \in C_c(G)$ we'll define $\|f\| = \max\{|f(x)| : x \in G\}$, where compactness of support insures the (finite) existence of the maximum.

Small and large functions. Since G is locally compact it has, at each point, a base of compact neighborhoods. In particular, there exists $r > 0$ such that $B_r(e)$ has compact closure. Thus $B_r(x) = xB_r(e)$ has compact closure for each $x \in G$. Fix this radius r for the rest of the proof.

Let $C_c^+(G)$ denote the collection of non-negative functions in $C_c(G)$. To say that $f \in C_c^+(G)$ is:

– *Small* means that its values are all ≤ 1 on G and its support lies in $B_{r/2}(x)$ for some $x \in G$.
– *Large* means that its values are all ≥ 1 on $B_r(x)$ for some $x \in G$.

Quasimeans. We'll call a positive linear functional on $C_c(G)$ a *quasimean* if it takes values ≤ 1 on small functions in $C_c^+(G)$ and ≥ 1 on large ones. Let $\mathscr{Q}$ denote the collection of quasimeans. We'll prove the existence of Haar measure for G by showing that $\mathscr{Q}$ is a nonempty, convex, weak-star compact subset of $C_c(G)^\sharp$ that is taken into itself by each translation-adjoint $L_\gamma^\sharp$. The usual argument involving the (solvable) Markov–Kakutani Theorem and the Riesz Representation Theorem will then lead to the desired Haar measure.

It's easy to check that $\mathscr{Q}$ is convex and, thanks to the fact that $\gamma B_r(x) = B_r(\gamma x)$, is also invariant under $L_\gamma^\sharp$ for each $\gamma \in G$.

$\mathscr{Q}$ is weak-star closed. Suppose $\Lambda \in C_c(G)^\sharp$ is a weak-star limit point of $\mathscr{Q}$. We wish to show that $\Lambda \in \mathscr{Q}$. For each $\varepsilon > 0$ and finite subset F of $C_c(G)$, the weak-star neighborhood of Λ

$$N(\Lambda, F, \varepsilon) = \{\Gamma \in C_c(G)^\sharp : |\Gamma(f) - \Lambda(f)| < \varepsilon \ \ \forall f \in F\}$$

contains a point of $\mathscr{Q}$. Suppose, then, that $f,g \in C_c^+(G)$ with f small and g large. Fix $\varepsilon > 0$, and choose $\Gamma \in N(\Lambda, \{f,g\}, \varepsilon) \cap \mathscr{Q}$. Then

$$\Lambda(f) \leq \Gamma(f) + \varepsilon \leq 1 + \varepsilon \qquad \text{and} \qquad \Lambda(g) \geq \Gamma(g) - \varepsilon \geq 1 - \varepsilon.$$

Since ε is an arbitrary positive number, $\Lambda(f) \leq 1$ and $\Lambda(g) \geq 1$, hence $\Lambda \in \mathscr{Q}$, as desired.

$\mathscr{Q}$ is weak-star compact. Since we now know that $\mathscr{Q}$ is weak-star closed, to show it's compact we need only prove that it's pointwise bounded on $C_c(G)$ (Corollary 9.15, p. 112). To this end fix $\Lambda \in \mathscr{G}$ and note that, thanks to the definition of "small" function: for every $f \in C_c^+(G)$ with support contained in some ball of radius $r/2$ we have $\Lambda(f) \leq \|f\|$. Now for arbitrary $f \in C_c^+(G)$ we can cover its (compact) support by a finite number of open d-balls of radius $r/2$. Lemma B.6 (p. 190) provides a partition of unity

$\{p_1, p_2, \ldots, p_n\}$ subordinate to that cover. Thus $f = \sum_{j=1}^n p_j f$ where each function $p_j f$ belongs to $C_c^+(G)$ and has support contained in a ball of radius $r/2$. It follows that

$$\Lambda(f) = \sum_{j=1}^n \Lambda(p_j f) \leq \sum_{j=1}^n \|p_j f\| \leq n\|f\|,$$

where the integer n depends on f, but not on Λ. Thus $\mathscr{Q}$ is pointwise bounded on $C_c^+(G)$.

Now suppose $f \in C_c(G)$. Then $f = f_+ - f_-$, the difference of two functions in $C_c^+(G)$, each of which has norm $\leq \|f\|$ and support contained in that of f. Thus

$$|\Lambda(f)| \leq \Lambda(f_+) + \Lambda(f_-) \leq n\|f_+\| + n\|f_-\| \leq 2n\|f\|,$$

where n does not depend on Λ. Thus $\mathscr{Q}$ is pointwise bounded on $C_c(G)$, hence weak-star compact in $C_c(G)^\sharp$.

$\mathscr{Q}$ is nonempty. For most proofs this sort of statement is a triviality. Not so here! None of the "usual suspects" (the point evaluations) belong to $\mathscr{Q}$. (Exercise: Why not?) What's needed is a subset S of G having the following properties:

(S1) S has at least one point in each open d-ball of radius r, and
(S2) S has no more than one point in each open ball of radius $r/2$.

Example 12.19. $G = \mathbb{R}$ and $r = 1$. Then $S = \mathbb{Z}$ has the desired properties. Note: S is *maximal* with respect to the property that any pair of its distinct elements lies at least 1 unit apart.

We're going to show that such a set S exists in every group G of the sort we're considering. Assuming this for the moment, define the functional Λ on $C_c(G)$ by

$$\Lambda(f) := \sum_{s \in S} f(s) \qquad (f \in C_c(G)). \tag{12.10}$$

Since we can cover each compact subset of G by finitely many open balls of radius $r/2$, each such set can contain at most finitely many points of S, hence for each

$f \in C_c(G)$ only finitely many summands on the right-hand side of (12.10) are nonzero. The right-hand side of (12.10) therefore makes sense and provides a positive linear functional on $C_c(G)$.

Suppose $f \in C_c^+(G)$ is "small," i.e., has support in a ball of radius $r/2$ and all values ≤ 1. Each such ball contains no more than one point of S, so $\Lambda(f) = 0$ or 1. Thus $\Lambda(f) \leq 1$ on "small" functions in $C_c^+(G)$. Suppose on the other hand that $f \in C_c^+(G)$ is "large," i.e., takes only values ≥ 1 on some ball of radius r. Then $\Lambda(f) \geq 1$ since this ball must contain a point of S. Thus $\Lambda \in \mathscr{Q}$, proving that $\mathscr{Q}$ is not empty.

The proof that G harbors the desired set S is inspired by Example 12.19.

Claim. Suppose $S \subset G$ is maximal with respect to the property

$$s,t \in S \text{ with } s \neq t \implies d(s,t) \geq r. \tag{*}$$

Then S satisfies conditions (S1) and (S2) above.

Proof of Claim. To check S has property (S1), note that if this were not the case there would be a point $x \in G$ at distance $\geq r$ from each point of S. Then $S \cup \{x\}$, which properly contains S, would obey (*) thus contradicting the maximality of S. As for (S2), suppose $s,t \in S$ lie in the ball $B_{r/2}(x)$. Then by the triangle inequality $d(s,t) < r$, hence $s = t$, thus establishing the Claim.

It remains to prove the existence of our maximal S. Let $\mathscr{T}$ denote the family of subsets T of G with the property (*).

$\mathscr{T}$ is nonempty. Since we're assuming the closure of $B_r(e)$ is compact, while G is not, there must exist $x \in G$ with $d(e,x) > r$. Thus $\{e,x\} \in \mathscr{T}$, so the family $\mathscr{T}$ is nonempty.

Enter Zorn's Lemma. If $\mathscr{C}$ is a subfamily of $\mathscr{T}$ that is *totally ordered* by inclusion (i.e., given two members of $\mathscr{C}$, one of them is contained in the other), then the union of the sets that are elements of $\mathscr{C}$ belongs to $\mathscr{T}$ (exercise). Thus each subfamily of $\mathscr{T}$ that is totally ordered has an upper bound, so by Zorn's Lemma (Appendix E.3) $\mathscr{T}$ has a maximal element S (note that this argument used only the fact that G is a metric space and $r < \sup_{x,y \in G} d(x,y)$).

Concluding the proof. We now have $\mathscr{Q}$, our nonempty, convex, weak-star compact subset of $C_c(G)^\sharp$, and the family $\mathscr{L}^\sharp$ of continuous affine (in fact linear) self-maps $L_\gamma^\sharp$ of $\mathscr{Q}$ ($\gamma \in G$). The argument in the paragraph preceding Theorem 12.6 shows that the family of maps $\mathscr{L}^\sharp$ inherits the solvability of G, hence our extended Markov–Kakutani Theorem (Theorem 12.1) guarantees that $\mathscr{L}^\sharp$ has a fixed point Λ in $\mathscr{Q}$. The measure provided for Λ by Riesz Representation Theorem is, by a familiar argument, the Haar measure we seek. □

Notes

Solvable families of maps. The original source for this is Day's paper [27, p. 285]; see also [37, Theorem 3.2.1, pp. 155–156].

The symmetric group. Exercise 12.4 is from [108, p. 253]. The fact that S_n is not solvable for $n \geq 5$ is a crucial step in the proof of Abel's Theorem: *For each $n \geq 5$ there is a polynomial of degree n whose roots can not all be found by radicals.* See, for example, Hadlock's Carus Monograph [43, Chap. 3].

Corollary 12.8. The amenability of solvable groups is due to von Neumann [88].

Right vs. left. Theorem 12.16, showing that each left-invariant mean gives rise to a bi-invariant one, is due to M.M. Day [27, Lemma 7, p. 285].

von Neumann's Paradox. The original source is [88]. For a modern exposition in English, see [121, Theorem 7.3, p. 99].

Exercise 12.10. This is taken directly from [102, Chap. 2, Exercise 17, p. 59].

Haar measure for solvable locally compact groups: Those simplifying assumptions. Every metrizable topological group has an invariant metric. In fact the *Birkhoff–Kakutani Theorem* asserts that every *first countable* group is metrizable and has such a metric. See [29, Corollary 3.10, p. 53] or [80, Sect. 1.22, pp. 34–36] for a proof, and [13, 57] for the original papers. With a bit more care the entire proof given above for the existence of Haar measure can be carried out for *every* solvable locally compact group. See Izzo [54] for how to do this in the commutative case; the solvable one being no different. The argument given above is just a translation of Izzo's proof to the solvable, invariantly metrizable case.