

# Chapter 9

## The Markov–Kakutani Theorem

FIXED POINTS FOR COMMUTING FAMILIES OF AFFINE MAPS

**Overview.** Consider the unit circle, the set  $\mathbb{T}$  of complex numbers of modulus one. Complex multiplication makes  $\mathbb{T}$  into a group, and the topology inherited from the complex plane makes it into a compact metric space. Here topology and algebra complement each other in that the group operations of multiplication  $\mathbb{T} \times \mathbb{T} \rightarrow \mathbb{T}$  and inversion  $\mathbb{T} \rightarrow \mathbb{T}$  are continuous. Tied up with the topology and algebra of  $\mathbb{T}$  is arc-length measure defined on the Borel subsets of  $\mathbb{T}$ , the salient property of which is its *rotation invariance*:  $\sigma(\gamma E) = \sigma(E)$  for each  $\gamma \in \mathbb{T}$  and Borel subset  $E$  of  $\mathbb{T}$ .

In this chapter we'll study a remarkable fixed-point theorem due to Markov and Kakutani, based on which we'll show that not just the unit circle, but in fact *every* compact abelian group, has such a “Haar measure”: a finite regular Borel probability measure invariant under the action of the group.<sup>1</sup> More generally, thanks again to the Markov–Kakutani theorem, we'll be able to produce both finitely and countably additive set functions that are invariant under quite general families of commuting transformations, a phenomenon that will point the way to our study in Chaps. 10–12 of the concepts of “amenability,” “solvability,” and “paradoxicality.”

**Prerequisites.** Some general topology: bases, compactness, product topologies, continuity of mappings. Basic measure theory. Acquaintance with (or at least willingness to believe) the Tychonoff Product Theorem and the version of the Riesz Representation Theorem that produces measures from positive linear functionals.

### 9.1 Topological groups and Haar measure

**Topological Groups.** Suppose  $G$  is a group with its operation written multiplicatively. We'll think of group multiplication as a map  $(x, y) \rightarrow xy$  that takes  $G \times G$  into  $G$ , and inversion  $x \rightarrow x^{-1}$  as a mapping of  $G$  into itself. If  $G$  has a topology (here,

<sup>1</sup> Haar measure is named for the Hungarian mathematician Alfred Haar (1885–1933). For further background see the *Notes* at the end of Chap. 13.

always Hausdorff) that renders these two maps continuous, we'll call  $G$ , endowed with this topology, a *topological group*. Thus the circle group  $\mathbb{T}$  described above is a compact topological group, and same is true of every product—both algebraic and topological—of  $\mathbb{T}$  with itself. Euclidean space  $\mathbb{R}^N$  with the usual topology and addition as its operation is a topological group that is not compact. Every group is a topological group in the discrete topology, the compact “discrete groups” being just the finite ones.

*Exercise 9.1.* Prove that:

- The unit circle  $\mathbb{T}$ , as described above, is a topological group.
- For each integer  $N \geq 2$  the product space  $\mathbb{T}^N$ , consisting of  $N$ -tuples of elements of  $\mathbb{T}$  is, with coordinatewise multiplication and the product topology (i.e., the topology it inherits from  $\mathbb{C}^N$ ), a compact topological group.
- $N$ -dimensional Euclidean space  $\mathbb{R}^N$  is a topological group with its usual topology and the operation of vector addition.
- $GL_N(\mathbb{R})$ , the collection of invertible  $N \times N$  real matrices, endowed with the usual matrix operations and the topology it inherits as a subset of  $\mathbb{R}^{N^2}$ , is a (non-commutative) topological group.

The most commonly studied topological groups are the *locally compact* ones, i.e., those for which at every point the topology has a base of compact neighborhoods. All the examples in Exercise 9.1, indeed all the groups we'll study from now on, are locally compact. Except for occasional digressions, we'll focus our attention on the compact ones.

*Exercise 9.2.* Show that every infinite subgroup of the circle group  $\mathbb{T}$  is dense. Use this result to show that the set of points  $\{\sin n : n \in \mathbb{Z}\}$  is dense in the closed unit interval.

**Borel sets and measures.** In a topological space the collection of *Borel sets* is the sigma algebra generated by the open sets. Since sigma algebras are closed under the taking of complements and countable unions, each closed subset is a Borel set, as are countable unions and intersections of Borel sets.

*Exercise 9.3* (Borel sets and continuity). Show that every continuous real-valued function on a topological space is measurable with respect to the Borel subsets of that space. Show that, at least for metric spaces, the sigma algebra of Borel sets is the smallest one with this property. Can you generalize this result beyond metric spaces?<sup>2</sup>

A *Borel measure* is simply a measure on the Borel sets of a topological space. To say a Borel measure is *regular* means that for every Borel set  $E$ :

$$\mu(E) = \inf\{\mu(U) : U^{\text{open}} \supset E\} = \sup\{\mu(K) : K^{\text{compact}} \subset E\} \quad (9.1)$$

i.e., the measure of each Borel set can be approximated arbitrarily closely from the outside by open sets and from the inside by compact ones.

---

<sup>2</sup> For more on this see the Notes at the end of this chapter.

In this chapter we'll consider only regular Borel measures that are positive and have total mass one, i.e., *regular Borel probability measures* (henceforth: RBPMs).

**Definition 9.1.** A *Haar measure* for a compact topological group  $G$  is an RBPM that is *invariant* under the group action in the sense that  $\mu(gB) = \mu(B)$  for every  $g \in G$  and Borel subset  $B$  of  $G$  (here  $gB$  is the set of elements  $gb$  as  $b$  runs through  $B$ ).

It turns out that every compact group has a (unique) Haar measure. In this chapter and the following two we'll use fixed-point theorems to prove this, concentrating for simplicity on the metrizable case. We'll discuss how these arguments can be enhanced to work in the general case, and in Chap. 12 will discuss an extension to locally compact groups.

*Some examples of Haar measure.* Arc-length measure (divided by  $2\pi$ ) for the unit circle  $\mathbb{T}$ , the product of arc-length measure (over  $2\pi$ ) with itself  $N$  times on  $\mathbb{T}^N$ , Lebesgue measure on  $\mathbb{R}^N$ .

*Exercise 9.4.* Show that (commutative or not) every *finite group*, in its discrete topology, has a unique Haar measure.

*Exercise 9.5.* Suppose  $G$  is a metrizable compact group with Haar measure  $\mu$ . Show that if  $E$  is a Borel subset of  $G$  with  $\mu(E) > 0$  then  $E \cdot E^{-1}$  (the set of points  $xy^{-1}$  with  $x$  and  $y$  in  $E$ ) contains an open ball.

*Suggestion:* Show that the function  $F: G \rightarrow [0, 1]$  defined by

$$F(x) = \int_G \chi_E(x^{-1}t)\chi_E(t) d\mu(t) \quad (x \in G)$$

is continuous on  $G$  and not identically zero (the metrizability of  $G$  is not really needed; it's there to simplify the proof of continuity for the integral).

*Left vs. right Haar measure.* For non-commutative compact groups what we've been calling Haar measure should more accurately be called "left Haar measure," to distinguish it from "right Haar measure," i.e., a regular Borel probability measure  $\mu$  for which  $\mu(Bg) = \mu(B)$  for each Borel set  $B$  and group element  $g$ . We'll see in Chap. 12 (Theorem 12.15) that for compact groups the two concepts are the same and that Haar measure is *unique*, but that the situation for non-compact groups is more complicated; see Exercise 12.6.

## 9.2 Haar Measure as a Fixed Point

**Measures and Functionals.** To each finite regular Borel measure  $\mu$  on a compact Hausdorff space  $Q$  there is an associated linear functional  $\Lambda_\mu$  defined on  $C(Q)$  (the space of continuous, real-valued functions on  $Q$ ) by

$$\Lambda_\mu(f) = \int f d\mu \quad (f \in C(Q)).$$

If  $\mu$  is a positive measure then the linear functional  $\Lambda_\mu$  is *positive*: it takes non-negative values on functions having only non-negative values. Everything we do from now on will depend upon the following famous result, which asserts that such  $\Lambda_\mu$ 's are the *only* positive linear functionals on  $C(Q)$ .

**The Riesz Representation Theorem for Compact Spaces.**<sup>3</sup> *If  $Q$  is a compact topological space and  $\Lambda$  is a positive linear functional on  $C(Q)$  then there is a unique positive regular finite Borel measure  $\mu$  on  $Q$  such that  $\Lambda = \Lambda_\mu$ .*

Regularity is important here; If  $Q$  is a nasty enough compact space, a positive linear functional on  $C(Q)$  may also be represented by a non-regular Borel probability measure (see, for example, [101, Chap. 2, Exercise 18, p. 59]). The good news: as shown by the exercise below, this can't happen for the most commonly occurring compact spaces.

*Exercise 9.6.* Show that for a compact metric space, every finite, positive Borel measure is regular.

*Suggestion:* Show that for such a measure  $\mu$ , the collection of subsets that satisfy condition (9.1) above (i.e., the  $\mu$ -regular sets) form a sigma algebra that contains all the closed sets.

**Invariance via Functionals.** For a compact topological group  $G$  (not necessarily commutative) and an RBPM  $\mu$  on the Borel subsets of  $G$ , what property of  $\Lambda_\mu$  corresponds to (left)  $G$ -invariance for  $\mu$ ?

Suppose  $\mu$  is an RBPM for  $G$ . Then by the change-of-variable formula of measure theory:

$$\int f(\gamma x) d\mu(x) = \int f(x) d\mu\gamma^{-1}(x) \quad (\gamma \in G, f \in C(G)), \quad (9.2)$$

where  $\mu\gamma^{-1}$  is the measure that gives the value  $\mu(\gamma^{-1}E)$  to the Borel subset  $E$  of  $G$ . Since  $G$ -invariance for  $\mu$  just means that  $\mu = \mu\gamma^{-1}$  for each  $\gamma \in G$ , Eq. (9.2) asserts that  $\mu$  is  $G$ -invariant if and only if

$$\int f(\gamma x) d\mu(x) = \int f(x) d\mu(x) \quad (9.3)$$

for every  $f \in C(G)$  and  $\gamma \in G$ . In order to rephrase this formula in terms of the linear functional  $\Lambda_\mu$ , let's define for each  $\gamma \in G$  the linear transformation  $L_\gamma: C(G) \rightarrow C(G)$  of (left) *translation* by  $\gamma$ :

$$(L_\gamma f)(x) = f(\gamma x) \quad (f \in C(G)). \quad (9.4)$$

In terms of the maps  $L_\gamma$ , the change-of-variable formula (9.2) becomes

$$\Lambda_\mu \circ L_\gamma = \Lambda_{\mu\gamma^{-1}} \quad (\gamma \in G) \quad (9.5)$$

---

<sup>3</sup> See [101, Theorem 2.14, pp. 40–41], where the theorem is proved for *locally compact* spaces.

for each RBPM  $\mu$  for  $G$ , while the invariance characterization (9.3) emerges as

$$\Lambda_\mu \circ L_\gamma = \Lambda_\mu \quad (\gamma \in G). \tag{9.6}$$

With these observations we’re one step away from being able to express an invariant measure as a fixed point. Here’s the step.

**Definition 9.2** (Dual space, adjoint). Let  $V$  be a real vector space and  $T : V \rightarrow V$  a linear transformation.

- (a) Denote by  $V^\sharp$  the *algebraic dual* of  $V$ , i.e., the vector space of all linear functionals (linear transformations  $V \rightarrow \mathbb{R}$ ) on  $V$ .
- (b) Define the *adjoint*  $T^\sharp$  of  $T$  by:  $T^\sharp \Lambda = \Lambda \circ T$  for  $\Lambda \in V^\sharp$ .

One checks easily that  $T^\sharp$  is a linear transformation  $V^\sharp \rightarrow V^\sharp$ . With these definitions the general transformation formula (9.5) becomes

$$L_\gamma^\sharp \Lambda_\mu = \Lambda_{\mu\gamma^{-1}} \quad (\gamma \in G), \tag{9.7}$$

while the invariance condition (9.6) can be written

$$L_\gamma^\sharp \Lambda_\mu = \Lambda_\mu \quad (\gamma \in G). \tag{9.8}$$

In summary:

**Proposition 9.3.** *An RBPM  $\mu$  on a compact group  $G$  is (left)  $G$ -invariant if and only if its associated linear functional  $\Lambda_\mu$  is a fixed point for each left-translation adjoint operator  $L_\gamma^\sharp : C(G)^\sharp \rightarrow C(G)^\sharp$  ( $\gamma \in G$ ).*

### 9.3 The Markov–Kakutani Fixed-Point Theorem

Having translated the problem of finding Haar measure for a compact group into one of finding a fixed point for a family of linear maps, let’s now turn our attention to a theorem that will guarantee the existence of such a fixed point. It turns out that some seemingly severe restrictions have to be made.

**Commutativity.** Our discussion of Haar measure began with the family of left-translation maps acting on the vector space  $C(G)$  of continuous real-valued functions on the compact group  $G$ , then moved on to the family of adjoints of these maps acting on the algebraic dual space  $C(G)^\sharp$ . If  $G$  is commutative then it’s easy to check that both families of maps—the translations and their adjoints—inheriting (under composition) the commutativity of  $G$ . Now commutativity is a natural condition to impose upon a family of maps for which one hopes to find a common fixed point; it’s an easy exercise to check that if a family of self-maps of some set commutes, then the set of fixed points of each map gets taken into itself by all the others. In

particular, if one of the maps has a *unique* fixed point (e.g., if it's a strict contraction of a complete metric space) then that's a common fixed point for the whole family.

However, as the example below shows, a commutative family of maps, each of which has a fixed point, need not have a common fixed point—even if the maps are all continuous on a compact metric space.

*Example 9.4.* Let  $S = \{1, 2, 3, 4, 5\}$  and  $\Phi = \{\varphi, \psi\}$  where  $\varphi$  fixes 3, 4, and 5, and interchanges 1 and 2, while  $\psi$  fixes 1 and 2, and takes 3 to 4, 4 to 5, and 5 to 3. In the notation and language of permutations:  $\varphi$  is the 2-cycle  $[1\ 2]$  (also called a “transposition”),  $\psi$  is the 3-cycle  $[3\ 4\ 5]$ , and being “disjoint” cycles,  $\varphi$  and  $\psi$  commute under composition. Thus  $S$  is compact in the discrete metric and  $\Phi$  is a commuting family of continuous maps, each of which has a fixed point but for which there is no common fixed point.

**Affine maps.** Example 9.4 above shows that for a family of self-maps of a topological space: continuity plus commutativity plus compactness is still not enough to insure a common fixed point. What extra condition can we add to remedy this situation? Recall that in Sect. 9.2 above we found that the problem of existence Haar measure on a compact group is equivalent to that of finding a common fixed point for a family of *linear* maps. It turns out that if we add to the hypotheses of continuity, compactness, and commutativity, additional conditions of convexity and “affine-ness” then common fixed points do exist.

**Definition 9.5** (Affine map). Suppose  $V$  is a real vector space,  $C$  a convex subset of  $V$ , and  $f$  is a map taking  $C$  into  $V$ . To say  $f$  is *affine* means that

$$f(tx + (1-t)y) = tf(x) + (1-t)f(y)$$

whenever  $x, y \in C$  and  $0 \leq t \leq 1$ .

Restrictions of linear maps to convex sets are affine; these are the only affine maps we'll consider here.

*Exercise 9.7.* Suppose  $V$  is a real vector space. Show that:

- If  $L$  is a linear map on the real vector space  $V$  and  $w$  is a vector in  $V$ , then the map  $v \rightarrow Lv + w$  is affine on  $V$ .
- The image of a convex subset of  $V$  under an affine map is again convex.
- Affine mappings of convex subsets  $C$  of  $V$  respect convex combinations, i.e., for all  $n$ -tuples of vectors  $(x_i : 1 \leq i \leq n)$  in  $C$  and non-negative scalars  $(t_i : 1 \leq i \leq n)$  that sum to 1,

$$f\left(\sum_{i=1}^n t_i x_i\right) = \sum_{i=1}^n t_i f(x_i).$$

**Vector Topology.** The algebraic setting for our fixed-point theorem will be quite restrictive: commutative families of affine maps. By contrast the *topological* setting will be very general: (real) *topological vector spaces*, i.e., vector spaces  $V$  over the real field on which there is a topology (which we'll always require to be Hausdorff)

that “respects” the vector operations. More precisely, the topology is required to render continuous<sup>4</sup>: addition, viewed as a map from the product space  $V \times V$  into  $V$ , and scalar multiplication, viewed as a map  $\mathbb{R} \times V \rightarrow V$ . Such a topology is called a *vector topology*. For example, the norm-induced topology of a normed linear space is a vector topology; we’ll soon discover others more suited to our purposes.

*Exercise 9.8.* Suppose  $U$  is a neighborhood of the zero vector in a topological vector space  $V$ . Show that  $V = \bigcup_{n \in \mathbb{N}} nU$ .

*Hint:* For each  $x \in V$  the map  $t \rightarrow tx$  ( $t \in \mathbb{R}$ ) takes the real line continuously into  $V$ .

With this foundation we’re now able to state the main result of this chapter.

**Theorem 9.6** (The Markov–Kakutani Theorem). *Suppose  $V$  is a topological vector space inside of which  $K$  is a nonvoid compact, convex subset. Suppose  $\mathcal{A}$  is a commutative family of continuous affine maps taking  $K$  into itself. Then there exists a point  $p \in K$  such that  $Ap = p$  for every  $A \in \mathcal{A}$ .*

Before proving the Markov–Kakutani Theorem, let’s sketch how it might be used to produce Haar measure for a compact abelian group  $G$ . Continuing the discussion of Sect. 9.2: the vector space  $V$  of the theorem will be the algebraic dual  $C(G)^\sharp$  of  $C(G)$  and the convex set  $K$  will be the set of linear functionals  $\Lambda_\mu$  on  $C(G)$ , where  $\mu$  runs through the collection of RBPMs on  $G$ . The family  $\mathcal{A}$  of affine self-maps of  $K$  will be the collection of *adjoints*  $L_\gamma^\sharp: C(G)^\sharp \rightarrow C(G)^\sharp$  of the translation operators  $L_\gamma$  for  $\gamma \in G$ . As mentioned earlier: it’s easily seen that  $\mathcal{A}$  inherits the commutativity of  $G$ .

Equation (9.7) guarantees that each of the maps  $L_\gamma^\sharp$  takes  $K$  into itself, so in order to apply the Markov–Kakutani theorem it remains to find a vector topology on  $V = C(G)^\sharp$  rendering  $K$  compact and each  $L_\gamma^\sharp$  continuous. Once this topology is found, the Markov–Kakutani Theorem will provide for  $\mathcal{A}$  a fixed point in  $K$  and, as pointed out in Sect. 9.2, the Riesz Representation Theorem will provide the  $G$ -invariant RBPM corresponding to this fixed point. All this we’ll do in Sect. 9.5. Right now, let’s prove the fixed-point theorem.

## 9.4 Proof of the Markov–Kakutani Theorem

We’ll break the proof into several pieces, the first being a straightforward consequence of the continuity of scalar multiplication. Throughout this section,  $V$  denotes a (real) topological vector space.

**Lemma 9.7.** *If  $K$  is a compact subset of  $V$  with  $0 \in K$ , then  $\bigcap_{n \in \mathbb{N}} n^{-1}K = \{0\}$ .*

*Proof.* Suppose  $U$  is a neighborhood of the zero vector in  $V$ . According to Exercise 9.8 above, the sets  $\{nU : n \in \mathbb{N}\}$  cover  $V$ , so they cover  $K$ . Since  $K$  is compact

---

<sup>4</sup> “Continuity” of a map in this context means that the inverse image of any open set is open.

there is a finite subcover. Since the sets  $nU$  increase with  $n$  there exists  $n \in \mathbb{N}$  such that  $K \subset nU$ , i.e.,  $n^{-1}K \subset U$ . Thus  $\bigcap_{n \in \mathbb{N}} n^{-1}K \subset U$  for each neighborhood  $U$  of zero in  $V$ . The desired result now follows from the fact that the topology of  $V$  is Hausdorff, so the intersection of all its zero-neighborhoods is  $\{0\}$ .  $\square$

The next result is the heart of the Markov–Kakutani Theorem: the special case where the commuting family  $\mathcal{A}$  consists of just a single map.

**Proposition 9.8.** *Suppose  $K$  is a compact, convex subset of  $V$  and  $A$  is an affine, continuous self-map of  $K$ . Then  $A$  has a fixed point in  $K$ . Moreover the set of all such fixed points is compact and convex.*

*Proof.* Let  $\mathbb{N}^* = \mathbb{N} \cup \{0\}$ , the set of non-negative integers. For  $n \in \mathbb{N}^*$  let  $A^n$  denote the composition of  $A$  with itself  $n$  times (with  $A^0$  denoting the identity map on  $K$ ). Then each map  $A^n$  is an affine, continuous self-map of  $K$ , as is each *arithmetic mean*  $M_n$  defined by

$$M_n x = \frac{1}{n+1} \sum_{j=0}^n A^j x \quad (x \in K, n \in \mathbb{N}^*).$$

Let  $S = \bigcap_{n \in \mathbb{N}^*} M_n(K)$ . Being an intersection of compact, convex sets,  $S$  is also compact and convex.

*Claim.*  $S$  is the fixed-point set of  $A$ .

*Proof of Claim.* Clearly every fixed point of  $A$  belongs to  $S$ . Conversely, fix  $y \in S$ . We wish to show that  $Ay = y$ . By the definition of  $S$ , for each  $n \in \mathbb{N}^*$  there is a vector  $x_n \in K$  such that  $y = M_n x_n$ . The map  $A$ , being affine, respects convex sums; in particular,  $AM_n = M_n A$  for each  $n$ . Thus

$$Ay - y = AM_n x_n - M_n x_n = M_n A x_n - M_n x_n,$$

i.e.,

$$Ay - y = \frac{1}{n+1} \sum_{j=0}^n (A^{j+1} x_n - A^j x_n) = \frac{1}{n+1} (A^{n+1} x_n - x_n) \in \frac{1}{n+1} (K - K),$$

where  $K - K$  is the set of all algebraic differences of pairs of elements of  $K$ . Since  $V$  is a topological vector space, the map  $V \times V \rightarrow V$  defined by  $(v, w) \rightarrow v + (-1)w$ , is continuous, so  $K - K$ , the image under this map of the compact set  $K$ , is compact. In the above calculation  $n$  is an arbitrary non-negative integer, so  $Ay - y$  belongs to  $\bigcap_{n \in \mathbb{N}^*} \frac{1}{n+1} (K - K)$  which, by Lemma 9.7 above (and the fact that  $0 \in K - K$ ), consists only of the zero vector. Thus  $Ay = y$ , as promised by the Claim.

*So Far.* We know that the compact, convex subset  $S = \bigcap_{n \in \mathbb{N}^*} M_n(K)$  of  $K$  is the fixed-point set of  $A$ .

*Remains to show.*  $S$  is nonempty. To this end, let  $\mathcal{M} = \{M_n : n \in \mathbb{N}^*\}$ , so  $\mathcal{M}(K) = \{M(K) : M \in \mathcal{M}\}$  is a family of closed subsets of the compact set  $K$ , with  $\bigcap \mathcal{M}(K) = S$ . If we can show that each finite subfamily of  $\mathcal{M}(K)$  has nonvoid

intersection, then by the finite intersection property of compact sets, the same will be true of  $\mathcal{M}(K)$  itself, thus finishing the proof.

Let  $\mathcal{F}$  be a finite subfamily of  $\mathcal{M}$  and let  $F$  be the composition of the maps  $\mathcal{F}$ , each map occurring exactly once in the composition. Since all the maps in  $\mathcal{M}$  commute under composition (exercise), in the definition of  $F$  they can occur in any order. Thus for each  $M \in \mathcal{F}$  we have  $F = M \circ H$  where  $H$  is a self-map of  $K$ , hence  $M(K) \supset M(H(K)) = F(K)$ . Conclusion:  $\bigcap_{M \in \mathcal{F}} M(K) \supset F(K) \neq \emptyset$ .  $\square$

Finally, we complete the proof of (the full-strength version of) the Markov–Kakutani Theorem.

**Proof of Theorem 9.6.** We’re given a compact, convex subset  $K$  of the topological vector space  $V$ , and a family  $\mathcal{A}$  of affine, continuous self-maps of  $K$  that commute under composition. Our goal is to show that there is a common fixed point for all the maps in  $\mathcal{A}$ .

For  $A \in \mathcal{A}$  let  $S_A = \{x \in K : Ax = x\}$ , the fixed-point set of  $A$ . From Proposition 9.8 we know that  $S_A$  is a convex, compact subset of  $K$  that is not empty. We desire to show that  $\bigcap_{A \in \mathcal{A}} S_A$ , the common fixed-point set for  $\mathcal{A}$ , is nonempty. For this it’s enough—again by the finite intersection property of compact sets—to show that

$$\bigcap_{A \in \mathcal{F}} S_A \neq \emptyset \tag{*}$$

for each finite subfamily  $\mathcal{F}$  of  $\mathcal{A}$ .

We proceed by induction on the number  $n$  of elements of  $\mathcal{F}$ , the case  $n = 1$  being just Proposition 9.8. Suppose (\*) is true for some  $n \geq 1$ , and that  $\mathcal{F}$  is a subfamily of  $\mathcal{A}$  consisting of  $n + 1$  maps. Pick a map  $A$  out of  $\mathcal{F}$  and let  $S$  denote the common fixed-point set of the  $n$  maps that remain. Then  $S$ , being the intersection of  $n$  compact, convex subsets of  $K$ , is again compact and convex in  $K$ ; by our induction hypothesis  $S \neq \emptyset$ . By commutativity,  $A(S_T) \subset S_T$  for each  $T \in \mathcal{F} \setminus \{A\}$ , hence  $A$  maps  $S$ , the intersection of these sets, into itself. By Proposition 9.8,  $A$  has a fixed point in  $S$ , which is therefore a common fixed point for  $\mathcal{F}$ . Conclusion: (\*) holds for each subfamily  $\mathcal{F}$  consisting of  $n + 1$  maps, so by induction it holds for every finite subfamily of  $\mathcal{A}$ .  $\square$

## 9.5 Markov–Kakutani Operating Manual

In order to enlist the Markov–Kakutani theorem in the production of invariant measures we need to find an appropriate vector topology for  $C(G)^\sharp$ , the algebraic dual space of  $C(G)$ . For this it’s best to think of  $C(G)$  as just a set, say  $S$ , and to view  $C(G)^\sharp$  as a subspace of  $\mathbb{R}^S$ , the vector space of *all* real-valued functions on  $S$ . It’s on this larger space that we’ll define our vector topology.

Let  $\omega(S)$  be the topology on  $\mathbb{R}^S$  for which each  $f \in \mathbb{R}^S$  has a base of neighborhoods defined as follows. For  $\varepsilon > 0$  and  $F$  a finite subset of  $S$ , let

$$N(f, F, \varepsilon) = \{g \in \mathbb{R}^S : |g(s) - f(s)| < \varepsilon \quad \forall s \in F\}. \quad (9.9)$$

The following exercise shows that the collection of sets (9.9) really is a base for a topology on  $\mathbb{R}^S$ .

*Exercise 9.9.* Show that if two sets  $N(f_j, F_j, \varepsilon_j)$  ( $j = 1, 2$ ) have nonempty intersection, then that intersection contains a third set  $N(f_3, F_3, \varepsilon_3)$ .

$\omega(S)$  is the *product topology* (general definition given in the paragraph after Exercise 9.11 below) one obtains by viewing  $\mathbb{R}^S$  as the topological product  $\prod_{s \in S} \mathbb{R}_s$ , where  $\mathbb{R}_s = \mathbb{R}$  for each  $s \in S$ . It is often called the “topology of pointwise convergence on  $S$ .” The next exercise explains why.

*Exercise 9.10.* Show that a sequence of real-valued functions on  $S$  converges in the topology  $\omega(S)$  if and only if it converges pointwise on  $S$ .

**Proposition 9.9.**  $\omega(S)$  is a vector topology on  $\mathbb{R}^S$ .

*Proof.* The first order of business is to show that the topology  $\omega(S)$  is *Hausdorff*. Given  $f_1$  and  $f_2$ , distinct functions in  $\mathbb{R}^S$ , we want to find neighborhoods  $N_j$  of  $f_j$  ( $j = 1, 2$ ) with  $N_1 \cap N_2 = \emptyset$ . Since  $f_1 \neq f_2$  there exists  $s \in S$  for which  $|f_1(s) - f_2(s)| = \varepsilon > 0$ . Then  $f_j \in N_j = N(f_j, \{s\}, \varepsilon/2)$  for ( $j = 1, 2$ ), and  $N_1 \cap N_2 = \emptyset$ .

It remains to establish *continuity* for the mappings  $\sigma : \mathbb{R}^S \times \mathbb{R}^S \rightarrow \mathbb{R}^S$  and  $\rho : \mathbb{R} \times \mathbb{R}^S \rightarrow \mathbb{R}^S$  of vector addition and scalar multiplication, defined respectively by

$$\sigma(f, g) = f + g \quad \text{and} \quad \rho(t, f) = tf \quad (f, g \in \mathbb{R}^S, t \in \mathbb{R}).$$

*Continuity of  $\sigma$ .* Suppose  $W$  is an open subset of  $\mathbb{R}^S$ . We need to show that  $\sigma^{-1}(W) = \{(f_1, f_2) \in \mathbb{R}^S \times \mathbb{R}^S : f_1 + f_2 \in W\}$  is open in  $\mathbb{R}^S \times \mathbb{R}^S$ . Fix  $(f_1, f_2) \in \sigma^{-1}(W)$  and choose  $\varepsilon > 0$  and a finite subset  $F$  of  $\mathbb{R}^S$  so that  $N(f_1 + f_2, F, \varepsilon) \subset W$ . Then  $U := N(f_1, F, \varepsilon/2) \times N(f_2, F, \varepsilon/2)$  is an open subset of  $\mathbb{R}^S \times \mathbb{R}^S$  that contains  $(f_1, f_2)$ . One checks easily that  $\sigma(U) \subset N(f_1 + f_2, F, \varepsilon) \subset W$ , hence  $U \subset \sigma^{-1}(W)$ . Thus  $\sigma^{-1}(W)$  is open in  $V$ , as desired.

*Continuity of  $\rho$ .* Fix  $f_0 \in \mathbb{R}^S$  and  $t_0 \in \mathbb{R}$ . Suppose we’re given  $\varepsilon > 0$  and a finite subset  $F$  of  $\mathbb{R}^S$ . Our goal is to find an open interval  $N_1$  about  $t_0$  and an  $\omega(S)$ -neighborhood  $N_2$  of  $f_0$  such that  $\rho(N_1 \times N_2) \subset N(t_0 f_0, F, \varepsilon)$ . In plain language, we’re looking for positive numbers  $\delta_1$  and  $\delta_2$ , and a finite subset of  $\mathbb{R}^S$ —which can only be  $F$  itself—such that:

$$|t - t_0| < \delta_1 \quad \text{and} \quad |f - f_0| < \delta_2 \quad \text{on } F \implies |tf - t_0 f_0| < \varepsilon \quad \text{on } F.$$

An “epsilon-halves” argument (exercise) shows that we can get the desired result by setting  $M = \max_{s \in F} |f_0(s)|$ , then taking  $\delta_1 = \frac{\varepsilon}{2M}$  and  $\delta_2 = \frac{\varepsilon}{2(|t_0| + \delta_1)}$ .  $\square$

*From Now on:* We’ll always assume  $\mathbb{R}^S$  to be endowed with the topology  $\omega(S)$ .

*Points as Functions.* We can view each point  $s$  of a set  $S$  as a function  $\hat{s} : \mathbb{R}^S \rightarrow \mathbb{R}$ , where  $\hat{s}(f) = f(s)$  for  $f \in \mathbb{R}^S$ . In the language of product spaces we can think of  $\hat{s}$  as the “projection” of  $\mathbb{R}^S$  onto its “ $s$ -th coordinate.”

**Proposition 9.10.** *For each  $s \in S$  the function  $\hat{s} : \mathbb{R}^S \rightarrow \mathbb{R}$  is continuous on  $\mathbb{R}^S$ ; the topology  $\omega(S)$  is the weakest one for which this is true.*

*Proof.* For  $t_0 \in \mathbb{R}$  and  $\varepsilon > 0$  let  $I$  be the open interval of radius  $\varepsilon$  centered at  $t_0$ . Then  $\hat{s}^{-1}(I) = N(f, \{s\}, \varepsilon)$  for all  $f \in \mathbb{R}^S$  for which  $f(s) = t_0$ . Thus the inverse image under  $\hat{s}$  of each real open interval is an open subset of  $\mathbb{R}^S$ , establishing the continuity of  $\hat{s}$ . Furthermore this argument shows that in every topology  $\tau$  on  $\mathbb{R}^S$  for which each of the functions  $\hat{s}$  is continuous,  $N(f, \{s\}, \varepsilon)$  has to be an open set, and since the basic open sets for  $\omega(S)$  are finite intersections of these, every  $\omega(S)$ -open set must be  $\tau$ -open, i.e., the topology  $\tau$  must be at least as strong as  $\omega(S)$ .  $\square$

**Compactness in  $\mathbb{R}^S$ .** The Markov–Kakutani Theorem requires compact sets. For finite dimensional normed linear spaces there are lots of these; the Heine–Borel Theorem asserts that every bounded subset of  $\mathbb{R}^N$  has compact closure. However we saw in Theorem 8.7 that nothing of the sort can happen once the dimension of our normed space becomes infinite. Fortunately, our vector topology  $\omega(S)$  on the space  $\mathbb{R}^S$  turns out to be weak enough to allow the re-emergence of Heine–Borel-like phenomena. The key to this is the *Tychonoff Product Theorem*, which states that arbitrary topological products of compact spaces are compact.<sup>5</sup> In its full generality Tychonoff’s theorem follows from the Axiom of Choice (Appendix E.3 below); for more on this see the *Notes* at the end of Appendix E.3.

**Definition 9.11.** To say a subset  $\mathcal{B}$  of  $\mathbb{R}^S$  is *pointwise bounded* means that for every  $s \in S$ ,  $\sup_{b \in \mathcal{B}} |b(s)| < \infty$  (i.e., the projection  $\hat{s}$  is bounded on  $\mathcal{B}$  for each  $s \in S$ ).

**Theorem 9.12** (A Heine–Borel Theorem for  $\mathbb{R}^S$ ). *Let  $S$  be a set and  $\mathcal{B}$  a subset of  $\mathbb{R}^S$ . Then  $\mathcal{B}$  has compact closure in  $\mathbb{R}^S$  if and only if it is pointwise bounded.*

*Proof.* (a) Suppose  $\mathcal{B}$  is pointwise bounded. For  $s \in S$  let  $m_s = \sup_{b \in \mathcal{B}} |b(s)|$ , and let  $I_s$  denote the compact real interval  $[-m_s, m_s]$ . Thus  $\mathcal{B}$  is a subset of the product space  $\mathcal{P} := \prod_{s \in S} I_s$ , which is compact by the Tychonoff Product Theorem. Now  $\mathcal{P}$  is a subset of  $\mathbb{R}^S$ ; it’s the collection of functions on  $f : S \rightarrow \mathbb{R}$  for which  $f(s) \in I_s$  for each  $s \in S$ . By its definition, the product topology on  $\mathcal{P}$  is just the restriction to that set of the topology  $\omega(S)$ . Since  $\mathcal{P}$  is compact in this topology, and  $\mathcal{B} \subset \mathcal{P}$ , the closure of  $\mathcal{B}$  lies in  $\mathcal{P}$  and so is compact.  
 (b) Suppose, conversely, that  $\mathcal{B}$  has compact closure in  $\mathbb{R}^S$ . Each “projection”  $\hat{s}$ , being continuous on  $\mathbb{R}^S$  (Proposition 9.10), is bounded on every compact subset, and in particular on  $\mathcal{B}$ , i.e.,  $\mathcal{B}$  is pointwise bounded.  $\square$

Suppose now that  $V$  is a real vector space. Note that  $V^\sharp$ , the algebraic dual of  $V$ , is a vector subspace of  $\mathbb{R}^V$ .

---

<sup>5</sup> See Exercise 9.11 below for an accessible special case of this.

**Definition 9.13.** (The Weak-Star Topology.) The restriction to  $V^\sharp$  of the product topology  $\omega(V)$  on  $\mathbb{R}^V$  is a vector topology that we'll call the *weak-star topology induced on  $V^\sharp$  by  $V$* .

The next result is crucial to the application of our infinite dimensional version of the Heine–Borel Theorem.

**Proposition 9.14.**  $V^\sharp$  is closed in  $\mathbb{R}^V$ .

*Proof.* We need to show that every limit point of  $V^\sharp$  in  $\mathbb{R}^V$  belongs to  $V^\sharp$ . So suppose  $\Lambda_0 \in \mathbb{R}^V$  is such a limit point; it's a real-valued function on  $V$  that we wish to prove is linear.

$\Lambda_0$  is additive. Fix  $x$  and  $y$  in  $V$ ; we wish to show that  $\Lambda_0(x+y) = \Lambda_0(x) + \Lambda_0(y)$ . To this end let  $\varepsilon > 0$  be given and consider the basic neighborhood  $U := N(\Lambda_0, \{x, y, x+y\}, \varepsilon/3)$  of  $\Lambda_0$ . Since  $\Lambda_0$  is a limit point of  $V^\sharp$  this neighborhood contains an element  $\Lambda$  of  $V^\sharp$ . By the definition of  $U$  (Eq. (9.9), p. 110) we have  $|\Lambda_0(w) - \Lambda(w)| < \varepsilon/3$  for  $w \in \{x, y, x+y\}$ . Thus

$$\begin{aligned} & |\Lambda_0(x+y) - \Lambda_0(x) - \Lambda_0(y)| \\ &= |\Lambda_0(x+y) - \Lambda(x+y) + [\Lambda(x) - \Lambda_0(x)] + [\Lambda(y) - \Lambda_0(y)]| \\ &\leq \underbrace{|\Lambda_0(x+y) - \Lambda(x+y)|}_{< \varepsilon/3} + \underbrace{|\Lambda(x) - \Lambda_0(x)|}_{< \varepsilon/3} + \underbrace{|\Lambda(y) - \Lambda_0(y)|}_{< \varepsilon/3} \\ &< \varepsilon. \end{aligned}$$

Since  $\varepsilon$  is an arbitrary positive number,  $\Lambda_0(x+y) - \Lambda_0(x) - \Lambda_0(y) = 0$ , as desired.

$\Lambda_0$  is homogeneous. Fix  $t \in \mathbb{R}$  and  $x \in V$ ; we wish to prove that  $\Lambda_0(tx) = t\Lambda_0(x)$ . Let  $\varepsilon > 0$  be given; set  $\delta = \varepsilon/2$  if  $|t| \leq 1$ , and  $= \varepsilon/(2|t|)$  otherwise. As before,  $N(\Lambda_0, \{x, tx\}, \delta)$  contains some  $\Lambda \in V^\sharp$ , so

$$\begin{aligned} |\Lambda_0(tx) - t\Lambda_0(x)| &= |\Lambda_0(tx) - \underbrace{\Lambda(tx) + t\Lambda(x) - t\Lambda_0(x)}_{=0}| \\ &\leq \underbrace{|\Lambda_0(tx) - \Lambda(tx)|}_{< \varepsilon/2} + \underbrace{|t| |\Lambda(x) - \Lambda_0(x)|}_{< \varepsilon/2} \\ &< \varepsilon. \end{aligned}$$

Thus (arbitrariness of  $\varepsilon$  once again)  $\Lambda_0(tx) - t\Lambda_0(x) = 0$ , as desired.  $\square$

**Corollary 9.15** (A Heine–Borel Theorem for Algebraic Duals). *For each real vector space  $V$ , a subset of  $V^\sharp$  is weak-star compact if and only if it is weak-star closed and pointwise bounded.*

*Proof.* Suppose  $\mathcal{E}$  is a weak-star compact subset of  $V^\sharp$ , hence closed therein. For each  $w \in V$  the coordinate projection  $\hat{w}: \Lambda \rightarrow \Lambda(w)$  is a continuous function  $V^\sharp \rightarrow \mathbb{R}$ , and so is bounded on  $\mathcal{E}$ . Thus  $\mathcal{E}$  is pointwise bounded on  $V$ .

Conversely, suppose  $\mathcal{E}$  is a weak-star closed subset of  $V^\sharp$  that is pointwise bounded on  $V$ . By Theorem 9.12,  $\mathcal{E}$  has compact closure in  $\mathbb{R}^V$ . By Proposition 9.14 the closure of  $\mathcal{E}$  in  $\mathbb{R}^V$  is the same as its closure in  $V^\sharp$ , which equals  $\mathcal{E}$ . Thus  $\mathcal{E}$  is weak-star compact in  $V^\sharp$ . □

*Exercise 9.11* (A “proto-Tychonoff” theorem). Suppose  $S$  is a countable set, say (without loss of generality)  $S = \mathbb{N}$ . Define a function  $\mathbb{R}^S \times \mathbb{R}^S \rightarrow [0, 1]$  by:

$$d(f, g) = \sum_{n=1}^{\infty} \frac{1}{2^n} \frac{|f(n) - g(n)|}{1 + |f(n) - g(n)|}.$$

- (a) Prove that  $d$  is a metric on  $\mathbb{R}^S$  and that  $\omega(S)$  is the topology it induces thereon.
- (b) Use sequential arguments to prove that  $\omega(S)$  is a vector topology on  $\mathbb{R}^S$ .
- (c) Use a diagonal argument to establish Theorem 9.12 for this special case.

*Exercise 9.12.* Suppose  $S$  is an uncountable set, and  $V$  is the set of functions in  $\mathbb{R}^S$  whose zero-set is at most countable. Show that  $V$  is a vector subspace of  $\mathbb{R}^S$  that is sequentially closed (i.e., every sequence in  $V$  that is  $\omega(S)$ -convergent has its limit in  $V$ ), but not closed. In particular, the topology  $\omega(S)$  is not metrizable.

*Recall our original motivation for the topology  $\omega(S)$ .* Given a compact abelian group  $G$ , we wished to apply the Markov–Kakutani Theorem to produce Haar measure for  $G$ . For this we needed to apply the theorem, not to the vector space  $C(G)$ , but rather to its algebraic dual space  $C(G)^\sharp$ . Thus we took the set  $S$  of the previous discussion to be  $C(G)$  itself, with the idea of restricting the topology  $\omega = \omega(C(G))$  to the dual space  $C(G)^\sharp$ . To complete our program we need to establish both the compactness of the convex set  $K$  of functionals  $\Lambda_\mu$  where  $\mu$  runs through the RBPMs on  $G$ , and the  $\omega$ -continuity of the commutative family  $\{L_\gamma^\sharp : \gamma \in G\}$  of adjoint self-maps of  $K$ .

*Clarity through abstraction:* Our arguments will be best understood in a more general setting. For this we’ll replace our compact abelian group  $G$  by a nonempty set  $S$ , assumed to carry no topology at all. We’ll replace the left-translation mappings furnished by the group operation with a commutative family  $\Phi$  of self-maps of  $S$ . Finally, the role of the space  $C(G)$  will be usurped by  $B(S)$ , the vector space of all functions  $f: S \rightarrow \mathbb{R}$  that are bounded, i.e., for which

$$\|f\| := \sup\{|f(s)| : s \in S\} < \infty. \tag{9.10}$$

It’s easy to check that  $\|\cdot\|$  is a norm that makes  $B(S)$  into a Banach space, but—perhaps surprisingly—we’ll never need this fact.

The vector space  $V$  to which we’ll apply the Markov–Kakutani Theorem will be  $B(S)^\sharp$ , endowed with the weak-star topology  $\omega$  it inherits as a subspace of  $\mathbb{R}^{B(S)}$ . The compact, convex subset  $K$  of the Markov–Kakutani Theorem will be the set of “means” in  $B(S)^\sharp$ , defined as follows:

**Definition 9.16.** A *mean* is an element of  $B(S)^\sharp$  that’s positive (takes non-negative values on non-negative functions) and takes the value 1 at the constant function 1.

*Notation.* We’ll use  $\mathcal{M}(S)$  to denote the collection of means in  $B(S)^\sharp$ .

*Exercise 9.13* (Evaluation functionals are means). For a set  $S$ :

- (a) Show that  $\mathcal{M}(S)$  is a convex subset of  $B(S)^\sharp$ .
- (b) Show that for each  $s \in S$  the *evaluation functional*  $f \rightarrow f(s)$  is a mean (so  $\mathcal{M}(S)$  is nonempty).

*Exercise 9.14* (Means and mean values). Each number claiming to be a “mean value” for a bounded function should at least lie between the function’s infimum and supremum. Show that  $\Lambda(f)$  has this property for each  $\Lambda \in \mathcal{M}(S)$  and  $f \in B(S)$ .

*Exercise 9.15.* Show that the convex hull of the evaluation functionals of Exercise 9.13(b)—a subset of  $\mathcal{M}(S)$  by that exercise—*exhausts*  $\mathcal{M}(S)$  if and only if  $S$  is a finite set.

Finally, the family of commuting affine maps for which we wish to find a common fixed point will be the adjoints of *composition operators* induced on  $B(S)$  by the maps in  $\Phi$ . More precisely, for each  $\varphi \in \Phi$  define  $C_\varphi : B(S) \rightarrow B(S)$  by  $C_\varphi f = f \circ \varphi$ . Clearly  $C_\varphi$  is a linear transformation  $B(S) \rightarrow B(S)$  that preserves positivity and fixes the constant functions. We’ll denote by  $\mathcal{C}_\Phi$  the collection of all these maps, and  $\mathcal{C}_\Phi^\sharp$  the collection of their *adjoints*. One checks easily that  $\mathcal{C}_\Phi^\sharp$  is commutative, and that each member of  $\mathcal{C}_\Phi^\sharp$  takes  $\mathcal{M}(S)$  into itself.

*So Far:* We’ve assembled the cast of characters demanded by the Markov–Kakutani Theorem, namely the vector space  $V = B(S)^\sharp$ , the commutative family of affine maps  $\mathcal{A} = \mathcal{C}_\Phi^\sharp = \{C_\varphi^\sharp : \varphi \in \Phi\}$ , the convex set  $K = \mathcal{M}(S)$  on which these maps act, and the vector topology  $\omega$  on  $B(S)^\sharp$  that’s going to glue these actors together. What’s left is to show that  $\mathcal{M}(S)$  is  $\omega$ -compact and each of the maps  $C_\varphi^\sharp$  is  $\omega$ -continuous.

**Corollary 9.17.**  $\mathcal{M}(S)$  is  $\omega$ -compact.

*Proof.* By Corollary 9.15 it’s enough to show that  $\mathcal{M}(S)$  is pointwise bounded on  $B(S)$  and  $\omega$ -closed in  $B(S)^\sharp$ . By Exercise 9.14 we know for each  $f \in B(S)$  and  $\Lambda \in \mathcal{M}(S)$  that the value  $\Lambda(f)$  lies between  $\inf_{s \in S} f(s)$  and  $\sup_{s \in S} f(s)$ . Thus  $\mathcal{M}(S)$  is a pointwise bounded subset of  $B(S)$ .

To show that  $\mathcal{M}(S)$  is  $\omega$ -closed in  $B(S)^\sharp$ , suppose  $\Lambda_0 \in B(S)^\sharp$  is a limit point of  $\mathcal{M}(S)$ . We wish to show that  $\Lambda_0 \in \mathcal{M}(S)$ , i.e., that  $\Lambda_0$ , which we already know is a linear functional on  $B(S)$ , is *positive* and “normalized” so that  $\Lambda_0(1) = 1$ .

To establish positivity, fix  $f \in B(S)$  with  $f \geq 0$  on  $S$ . Let  $\varepsilon > 0$  be given. Then the basic  $\omega$ -neighborhood  $N(\Lambda_0, \{f\}, \varepsilon)$  of  $\Lambda_0$  contains a point  $\Lambda \in \mathcal{M}(S)$ . Thus

$$\Lambda(f) - \Lambda_0(f) \leq |\Lambda(f) - \Lambda_0(f)| < \varepsilon, \quad (9.11)$$

hence  $\Lambda(f) \leq \Lambda_0(f) + \varepsilon$  for every  $\varepsilon > 0$  so  $\Lambda(f) \leq \Lambda_0(f)$ . But  $0 \leq \Lambda(f)$  since  $\Lambda$ , being a member of  $\mathcal{M}(S)$ , is a positive linear functional on  $B(S)$ . Conclusion:  $0 \leq \Lambda_0(f)$ , thus establishing the positivity of the limit functional  $\Lambda_0$ .

As for “normalization,” choose  $f \equiv 1$  on  $S$ . By the second inequality of (9.11): for each  $\varepsilon > 0$ ,  $|1 - \Lambda_0(f)| < \varepsilon$  hence  $\Lambda_0(f) = 1$ .  $\square$

It remains to establish  $\omega$ -continuity for the adjoint maps  $C_\phi^\sharp$ . This follows from something more general:

**Proposition 9.18** (Adjointness Are continuous). *If  $V$  is a real vector space and  $T$  a linear transformation of  $V$  into itself, then the adjoint map  $T^\sharp$  is weak-star continuous on  $V^\sharp$ .*

*Proof.* In a topological vector space  $V$ , the map  $x \rightarrow x + h$  of translation by a fixed vector  $h$  is a linear homeomorphism, so a linear transformation on  $V$  is continuous if and only if it is continuous at the origin. Thus it’s enough to show that  $T^\sharp$  is continuous at the origin of  $V^\sharp$ , and for this it’s enough to show that the  $T^\sharp$ -inverse image of each basic zero-neighborhood in  $V^\sharp$  contains a basic zero-neighborhood. For this, suppose  $\varepsilon > 0$  and  $F$  is a finite subset of  $V$ . Upon chasing definitions one sees that  $(T^\sharp)^{-1}(N(0, F, \varepsilon)) = N(0, T(F), \varepsilon)$ .  $\square$

Thus the Markov–Kakutani Theorem applies to the triple  $(B(S)^\sharp, \mathcal{M}(S), \mathcal{C}_\Phi^\sharp)$ ; it yields the following:

**Theorem 9.19** (Invariant means). *Suppose  $\Phi$  is a commutative family of self-maps of a set  $S$ . Then there is a mean  $\Lambda$  in  $B(S)^\sharp$  such that  $C_\phi^\sharp \Lambda = \Lambda$  (i.e.,  $\Lambda \circ \phi = \Lambda$ ).*

## 9.6 Invariant Measures for Commuting Families of Maps

Theorem 9.19 yields, as a special case, the theorem that started our quest.

**Corollary 9.20** (Haar Measure for compact abelian groups). *Every compact abelian group  $G$  supports, on its Borel subsets, a  $G$ -invariant RBPM.*

*Proof.* Apply Theorem 9.19 with  $S = G$  and  $\Phi$  the collection of translation-maps  $x \rightarrow \gamma \cdot x$  for  $\gamma$  and  $x$  in  $G$ . The resulting composition operators on  $B(G)$  are the translation operators  $L_\gamma$  on  $B(G)$ . For this situation Theorem 9.19 provides a mean  $\Lambda$  on  $B(G)$  that’s fixed by each of the transformations  $L_\gamma^\sharp$  for  $\gamma \in G$ , so by the Riesz Representation Theorem and Proposition 9.3, the restriction of this functional to  $C(G)$  has the form  $\Lambda_\mu$ , where  $\mu$  is an RBPM on the Borel subsets of  $G$ .  $\square$

The argument above gives a more general result:

**Corollary 9.21.** *Suppose  $Q$  is a compact Hausdorff space and  $\Phi$  is a commutative family of continuous self-maps of  $Q$ . Then there is an RBPM  $\mu$  on the Borel subsets of  $Q$  that is  $\Phi$ -invariant in the sense that for every  $f \in C(Q)$  and  $\phi \in \Phi$ ,*

$$\int f \circ \phi d\mu = \int f d\mu$$

or equivalently, for every Borel subset  $B$  of  $Q$ ,

$$\mu(\varphi^{-1}(B)) = \mu(B).$$

*Proof.* Exercise. □

*Example 9.22.* Let  $D$  denote the closed unit disc of  $\mathbb{R}^2$ , and  $\Phi$  the restrictions to  $D$  of rotations of  $\mathbb{R}^2$  about the origin. Then  $\Phi$  is a commutative family of maps, each of which takes  $D$  continuously onto itself. Thus Corollary 9.21 guarantees for  $D$  an RBPM invariant under each member of  $\Phi$ . In fact *two* such measures come immediately to mind: Lebesgue area measure on  $D$  and arc-length measure on  $\partial D$ , the unit circle (both measures normalized to have total mass one). Thus the uniqueness established above for Haar measure on compact abelian groups fails for the more general case of RBPMs invariant under commutative families of maps.

*Exercise 9.16.* Show that there is an uncountable family of RBPM's on  $D$  invariant under the rotation group  $\Phi$  defined above.

**Remark 9.23** (Role of commutativity). In the arguments above, the hypothesis of commutativity imposed upon the family of maps  $\Phi$  showed up only at the very end, where it legitimized our use of the Markov–Kakutani Theorem. In Chap. 12 we'll extend the Markov–Kakutani Theorem to families of maps that are “almost commutative,” (e.g., to groups of maps that are *solvable*). Here the argument that proved Theorem 9.19 will go through *verbatim*, with commutativity replaced at the final stage by the new hypothesis on the family  $\Phi$  of self-maps of  $S$ . As a corollary we'll obtain the existence of Haar measure for compact, solvable groups. To obtain Haar measure for *all* compact groups, however, will require a new fixed-point theorem; this we'll explore in Chap. 13.

## 9.7 Harmonic Analysis on Compact Abelian Groups

The existence of Haar measure for compact abelian groups allows us to generalize to that context the Fourier analysis that's so important for functions that are integrable on the unit circle  $\mathbb{T}$ . Recall that Haar measure  $\mu$  on  $\mathbb{T}$  is Lebesgue arc-length measure normalized to have unit mass. For  $1 \leq p \leq \infty$ , let's denote the (complex) Lebesgue space of  $\mathbb{T}$  with respect to this measure by  $L^p(\mathbb{T})$ . Let  $\gamma_n(\zeta) = \zeta^n$  for  $\zeta \in \mathbb{T}$ , and for  $f \in L^1(\mathbb{T})$  and  $n \in \mathbb{Z}$ , define the  $n$ -th Fourier coefficient of  $f$  by

$$\hat{f}(n) = \int_{\mathbb{T}} f \gamma_n^{-1} d\mu = \frac{1}{2\pi} \int_0^{2\pi} f(e^{i\theta}) e^{-in\theta} d\theta = \langle f, \gamma^n \rangle,$$

where  $d\mu(e^{i\theta}) = \frac{d\theta}{2\pi}$ .

Now  $L^2 = L^2(\mathbb{T})$  is a Hilbert space with inner product

$$\langle f, g \rangle = \int f \bar{g} d\mu \quad (f, g \in L^2)$$

so  $\hat{f} = \langle f, \gamma_n \rangle$  for  $f \in L^2$  and  $n \in \mathbb{Z}$ . It's easy to check that the exponential functions  $\{\gamma_n : n \in \mathbb{Z}\}$  form an orthonormal subset of  $L^2$ . The linear span  $\mathcal{S}$  of this orthonormal set (the set of *trigonometric polynomials*) is a subalgebra of  $C(\mathbb{T})$  that's closed under complex conjugation and separates points of  $\mathbb{T}$  (a feat accomplished single-handedly by  $\gamma_1$ , which is the identity map on  $\mathbb{T}$ ). Thus the Stone-Weierstrass Theorem (see, e.g., [101, Theorem 7.33, p. 165]) assures us that  $\mathcal{S}$  is dense in the max-norm topology of  $C(\mathbb{T})$ , and since  $C(\mathbb{T})$  is dense in  $L^2$ , and the  $L^2$ -topology is weaker than that of  $C(\mathbb{T})$ , we see that  $\mathcal{S}$  is dense in  $L^2$ .

*Conclusion:* The exponential functions  $\{\gamma_n : n \in \mathbb{Z}\}$  form an *orthonormal basis* for  $L^2$ , hence for every function  $f$  in that space we have

$$\|f\|_2^2 = \sum_{n \in \mathbb{Z}} |\hat{f}(n)|^2$$

from which it follows that

$$f = \sum_{n \in \mathbb{Z}} \langle f, \gamma_n \rangle \gamma_n = \sum_{n \in \mathbb{Z}} \hat{f}(n) \gamma_n \quad (9.12)$$

with the series convergent in  $L^2$  *unconditionally* in the sense that for every  $\varepsilon > 0$  there exists a finite subset  $F_\varepsilon$  of  $\mathbb{Z}$  such that

$$F \supset F_\varepsilon \implies \left\| \sum_{n \in F} \hat{f}(n) \gamma_n - f \right\| < \varepsilon.$$

The series in (9.12) is called the *Fourier series* of  $f$ ; it represents a decomposition of that function into “frequencies”  $\gamma_n$ .

All this is standard, and forms the basis for the harmonic analysis of square-integrable functions on the unit circle. In order to generalize this theory to other compact abelian groups, we need the following observation:

**Proposition 9.24.** *The exponential functions  $\{\gamma_n : n \in \mathbb{Z}\}$  are precisely the continuous homomorphisms of  $\mathbb{T}$  into itself.*

One checks easily that each function  $\gamma_n$  is indeed a homomorphism of  $\mathbb{T}$  (meaning:  $\gamma_n(\zeta \eta^{-1}) = \gamma_n(\zeta) \gamma_n(\eta)^{-1}$  for each  $\zeta, \eta \in \mathbb{T}$  and  $n \in \mathbb{Z}$ ). The exercise set below shows that the  $\gamma_n$ 's are the *only* continuous homomorphisms of  $\mathbb{T}$ .

*Exercise 9.17.* Suppose  $\Gamma : \mathbb{R} \rightarrow \mathbb{T}$  is a continuous group homomorphism, where  $\mathbb{R}$  has its additive structure. Thus  $\Gamma(0) = 1$ ,  $\Gamma(x+y) = \Gamma(x)\Gamma(y)$  and  $\Gamma(-x) = \Gamma(x)^{-1}$  for each  $x, y \in \mathbb{R}$ .

- Suppose in addition that  $\Gamma$  is *differentiable* at the origin. Show that  $\Gamma$  is differentiable at every  $x \in \mathbb{R}$ , with  $\Gamma'(x) = \Gamma'(0)\Gamma(x)$ . Conclude that  $\Gamma(x) = e^{i\lambda x}$  for each  $x \in \mathbb{R}$ , where  $\lambda := \Gamma'(0)$ .
- Not assuming the differentiability of  $\Gamma$ , show that there exists  $\delta \in (0, 2\pi)$  for which  $A := \int_0^\delta \Gamma(t) dt \neq 0$ . Show that this implies  $A\Gamma(x) = \int_{s=x}^{x+\delta} \Gamma(s) ds$ . Conclude that  $\Gamma$  is differentiable on  $\mathbb{R}$ .

- (c) Show that if  $\Gamma$  is  $2\pi$ -periodic on  $\mathbb{R}$  then the constant  $\lambda$  of part (a) must be an integer.  
 (d) Now use part (c) to finish the proof of Proposition 9.24.

**Definition 9.25.** A *character* of a topological group  $G$  is a continuous homomorphism  $G \rightarrow \mathbb{T}$ .

*Notation.* We'll use  $\hat{G}$  to denote the set of characters of  $G$ . Thus  $\hat{\mathbb{T}} = \{\gamma_n : n \in \mathbb{Z}\}$ .

*Exercise 9.18 (Dual Group).* Show that  $\hat{G}$ , with pointwise multiplication, is a group. It's called the *dual group* of  $G$ .

*Exercise 9.19 (Some Dual Groups).* For  $\lambda \in \mathbb{R}$  define the exponential function  $\gamma_\lambda(x) := e^{i\lambda x}$  for  $x \in \mathbb{R}$ . Think of  $\mathbb{R}$  as a group with its usual additive structure, and  $\mathbb{Z}$  as a subgroup of  $\mathbb{R}$ . Show that:

- (a)  $\hat{\mathbb{R}}$  is group-isomorphic to  $\mathbb{R}$  via the identification  $\lambda \rightarrow \gamma_\lambda$ ,  $\lambda \in \mathbb{R}$ .  
 (b)  $\hat{\mathbb{T}}$  is group-isomorphic to  $\mathbb{Z}$  via the identification  $n \rightarrow \gamma_n$ ,  $n \in \mathbb{Z}$ .

For a compact abelian group  $G$  it's common to denote by  $L^2(G)$  the  $L^2$ -space defined for the Haar measure of  $G$ .

*Exercise 9.20.* If  $G$  is a compact abelian group then  $\hat{G}$  is an orthonormal subset of  $L^2(G)$ .

Just as for the unit circle,  $\hat{G}$  is an orthonormal *basis* for  $L^2(G)$ . As before,  $\hat{G}$  is closed under complex conjugation (easy), and separates points of  $G$  (not easy: see, e.g., [100, Sect. 1.5.2, p. 24]), so its linear span—what we might think of as the collection of *trigonometric polynomials* on  $G$ —is, by Stone-Weierstrass, dense in  $C(G)$ , hence also in  $L^2(G)$ .

If  $f \in L^2(G)$  and  $\gamma \in G$ , we define the *Fourier transform*  $\hat{f} : \hat{G} \rightarrow \mathbb{C}$  by

$$\hat{f}(\gamma) := \int_G f\gamma^{-1} d\mu \quad (\gamma \in \hat{G}),$$

where  $\mu$  is Haar measure for  $G$ . Then, as in the circle case:

**Proposition 9.26.** *If  $G$  is a compact abelian group then for each  $f \in L^2(G)$ ,*

$$\sum_{\gamma \in \hat{G}} |\hat{f}(\gamma)|^2 = \|f\|_2^2 \quad \text{and} \quad \sum_{\gamma \in \hat{G}} \hat{f}(\gamma)\gamma = f.$$

These formulae employ the same kind of “unordered summation” used for Fourier series on the circle group, with the term “Fourier series” once again denoting the character series representing  $f$ .

*Remark.* The characters of *non-commutative* groups never separate points. Indeed, if  $g$  and  $h$  are non-commuting elements of a group  $G$ , and  $\gamma$  is a character on  $G$ , then  $\gamma(gh) = \gamma(g)\gamma(h) = \gamma(hg)$ . It gets worse; Exercise 13.8 (p. 178) shows that even for compact groups it's possible for the only character to be the trivial one  $\gamma \equiv 1$ !

## Notes

*Commuting families of maps.* Apropos to Example 9.4: A long-standing problem asked if every commuting family of continuous self-maps of the *closed unit interval* had to have a common fixed point. Counterexamples were published in 1969, independently by Boyce [15] and Hunecke [52]; their constructions are nontrivial.

*Exercise on Borel sets.* Thanks to Urysohn's Lemma the result of Exercise 9.3 remains true for normal (Hausdorff) topological spaces. However it is *not* true in general. Let  $X$  denote the space of ordinal numbers less than or equal to the first uncountable one  $\Omega$ , taken in the interval topology. Then  $X$  is a Hausdorff space, but each continuous real-valued function on  $X$  is constant on some final segment  $[\alpha, \Omega]$ ,  $\alpha$  a countable ordinal. From this it follows that, although the singleton  $\{\Omega\}$  is a Borel set (it is closed), it does not belong to  $\mathcal{C}$ , the smallest sigma algebra rendering each continuous real-valued function measurable. Thus  $\mathcal{C}$  is strictly smaller than the sigma algebra of Borel subsets of  $X$ .

*The Markov–Kakutani Theorem.* The proof of Theorem 9.6 is Kakutani's proof from [56]. Markov earlier gave a proof [75] for locally convex spaces, using Tychonoff's extension [120] to that setting of the Schauder Fixed-Point Theorem.

*The Tychonoff Product Theorem.* For a proof that the Axiom of Choice implies the Tychonoff Product Theorem see [103, Theorem A2, pp. 392–393], or [55, Sect. 2.2, p. 11] for a proof based on “filters,” or [25] for one based on “nets.”

The Tychonoff Product Theorem does not require that the factors of the product be Hausdorff; in fact this “non-Hausdorff” version of the theorem is actually *equivalent* to the Axiom of Choice (see, e.g., [55, Sect. 2.6, p. 26, Problem 8]).

*The Riesz Representation Theorem.* The first result of this type appeared in a 1909 paper of Frigyes Riesz [97], who proved that every continuous linear functional on the Banach space  $C([0, 1])$  is represented by Stieltjes integration against a real-valued function on  $[0, 1]$  of bounded variation. The “positive” version we've been using above follows easily from this one.