Chapter 13
Beyond Markov–Kakutani

THE RYLL–NARDZEWSKI FIXED-POINT THEOREM

Overview. In the last chapter we extended the Markov–Kakutani Theorem—originally proved only for *commuting* families of continuous affine maps—to "solvable" families of such maps. We used our enhanced theorem to show that every solvable group is amenable and that Haar measure exists for every topological group that is both solvable and compact. By contrast, we've seen (Chap. 11) that the group $\mathscr{R}$ of origin-centric rotations of $\mathbb{R}^3$ is paradoxical, hence not amenable, and therefore not solvable. Now $\mathscr{R}$ is naturally isomorphic to the group $\mathrm{SO}(3)$ of 3×3 orthogonal real matrices with determinant 1 (Appendix D), a group easily seen to be compact. Thus not every compact group is amenable.

Conclusion: Fixed-point theorems that produce invariant means cannot prove the existence of Haar measure for every compact group.

In this chapter we'll turn to a fixed-point theorem in which the Markov–Kakutani hypothesis of solvability is replaced by a topological condition of "uniform injectivity." This result, due to the Polish mathematician Czesław Ryll–Nardzewski, works with an appropriate modification of our previous duality method to provide Haar measure for *all* compact topological groups. For ease of exposition we'll focus on compact groups that are *metrizable,* sketching afterwards how to make the arguments work in general. Finally, we'll identify the Haar measure for $\mathrm{SO}(3)$.

Throughout this chapter we'll be working in vector spaces over the real numbers.

13.1 Introduction to the Ryll–Nardzewski Theorem

Theorem 13.1 (Ryll–Nardzewski [105])**.** *Suppose X is a* locally convex *topological vector space in which K is a nonvoid, compact, convex subset. Suppose $\mathscr{S}$ is a* semigroup *of continuous, affine self-maps of K that is* uniformly injective. *Then $\mathscr{S}$ has a fixed point in K.*

J.H. Shapiro, *A Fixed-Point Farrago*, Universitext,
DOI 10.1007/978-3-319-27978-7_13

We'll first seek to understand the hypotheses of the Ryll–Nardzewski Theorem after which we'll prove it for the *duals of separable Banach spaces,* a special case that's still general enough to provide the existence of Haar measure for compact, metrizable groups, and which gives an accurate guide to the proof of the general theorem. Let's start with the undefined terms in the Theorem's statement, taking them in the order in which they occur.

Semigroup. A set with an associative binary operation. If $\mathscr{F}$ is a family of self-maps of some set S, then the *semigroup generated by* $\mathscr{F}$ (its operation being composition of maps) consists of all possible finite compositions of maps in $\mathscr{F}$. This is the smallest semigroup of self-maps of S containing $\mathscr{F}$; it has the same set of common fixed points as $\mathscr{F}$, and one can even throw in the identity map without changing the common-fixed-point set. Thus when considering fixed points for families of self-maps one need only consider "compositional semigroups with identity."

Locally Convex. For a topological vector space (always assumed Hausdorff) this property means that each point has a base of convex neighborhoods. We have already worked with several important examples of locally convex spaces:

(a) Normed linear spaces: the balls centered at a given point are convex and form a base for the neighborhoods of that point.
(b) The space $\mathbb{R}^S$ of all real-valued functions on a set S, with its topology of "pointwise convergence": the basic neighborhoods $N(f,F,\varepsilon)$ for this topology, as defined by Eq. (9.9) (p. 110), are all convex.
(c) The weak-star topology induced on the algebraic dual $V^\sharp$ of a real or complex vector space V.

A version of the Hahn–Banach Theorem guarantees, for each locally convex topological vector space, the existence of enough continuous linear functionals to separate distinct points of the space, and more generally, to separate disjoint closed convex sets.[1] The exercise below shows that in the absence of local convexity such separation is not guaranteed.

Exercise 13.1 (Non locally convex pathology[2]). For $0 < p < 1$ consider the space $L^p = L^p([0,1])$ consisting of (a.e.-equivalence classes of) real-valued Lebesgue measurable functions f on the unit interval for which $\|f\| = \int_0^1 |f(x)|^p\,dx < \infty$ (omission of the p-th root of the integral on the right is deliberate). For $f,g \in L^p$ let $d(f,g) = \|f-g\|$. Show that:

(a) $\|\cdot\|$ is not a norm, but d is a metric making L^p into a topological vector space.
(b) On L^p the topology induced by the metric d is not locally convex. In fact, the only open (nonempty) convex set is the whole space!
(c) The only continuous linear functional on L^p is the zero-functional.

Uniformly Injective. "Injective" is another way of saying "one-to-one." To say a family of maps $\mathscr{F}$ taking a set S into a topological vector space X is *uniformly*

[1] See, e.g., [103, Chap. 3, pp. 56–62].
[2] For more details, see [103, Sect. 1.47, pp. 36–37].

injective[3] means that for each pair of distinct points $s,t \in S$ the zero vector does not belong to the closure of the set $\{f(s) - f(t) : f \in \mathscr{F}\}$. If X is a normed linear space, then uniform injectivity for $\mathscr{F}$ means that for every pair $s,t \in S$ with $s \neq t$ there exists a positive number $\delta = \delta(s,t)$ such that:

$$\delta < \|f(s) - f(t)\| \quad \text{for every} \quad f \in \mathscr{F}. \tag{13.1}$$

Why Injectivity? Let $X = \mathbb{R}$, $K = [0,1]$, and consider the two-element compositional semigroup $\mathscr{S} = \{\varphi, \psi\}$, where $\varphi \equiv 0$ and $\psi \equiv 1$. Thus X is locally convex, K is a nonempty, compact, convex subset of X, and $\mathscr{S}$ is a finite semigroup of affine, continuous self-maps of K that does not have a common fixed point. The following prototype of the Ryll–Nardzewski Theorem shows that the culprit here is "lack of injectivity."

Proposition 13.2. *Suppose K is a nonvoid, compact, convex subset of a topological vector space. Then every finite semigroup of continuous, injective, affine self-maps of K has a common fixed point.*

Proof. Let $\mathscr{S}$ denote our finite semigroup of maps. As noted earlier, there is no loss of generality in assuming that it contains the identity map e_K on K.

Claim. $\mathscr{S}$ is a group.

Proof of Claim. We need only show that each map in $\mathscr{S}$ has an inverse. Fix $A \in \mathscr{S}$ and note that since $\mathscr{S}$ is finite there exist positive integers n and m with $1 \leq m < n$ such that $A^n = A^m$, (where, e.g., A^n denotes the composition of A with itself n times). Thus $A^n = A^m A^{n-m} = A^n A^{n-m}$, and since A^n is injective

$$e_K = A^{n-m} = A^{n-m-1}A = AA^{n-m-1}.$$

This exhibits A^{n-m-1} (which exists and belongs to $\mathscr{S}$ because $n - m \geq 1$) as the compositional inverse of A, thus proving the Claim.

Having established that $\mathscr{S}$ is a group, let $A_1, A_2, \ldots, A_n$ denote its elements, and denote by A_0 the arithmetic mean of these elements:

$$A_0 x = \frac{1}{n} \sum_{j=1}^{n} A_j x \qquad (x \in K).$$

Now A_0, though perhaps not a member of $\mathscr{S}$, is nonetheless a continuous, affine self-map of K. The Markov–Kakutani Theorem[4] therefore guarantees that A_0 has a fixed point x_0 in K.

[3] Alternative terminology: "non-contracting," or in dynamical systems: "distal."

[4] For this we need only the "single-map" version: Proposition 9.8, p. 108.

Fix $A \in \mathscr{S}$. Since A is affine it respects convex combinations, so $AA_0 = (1/n)\sum_{j=1}^{n} AA_j$. Since $\mathscr{S}$ is a group, the n-tuple $(AA_1, AA_2, \dots, AA_n)$ is a permutation of the original list $(A_1, A_2, \dots, A_n)$ of the elements of $\mathscr{S}$. Conclusion: $AA_0 = A_0$. Consequently

$$x_0 = A_0 x_0 = AA_0 x_0 = Ax_0$$

i.e., x_0 is a fixed point for A, hence a common fixed point for $\mathscr{S}$. □

On the other hand, for *infinite* semigroups of affine, continuous maps: *injectivity alone is not enough to guarantee a common fixed point.* Once again let $X = \mathbb{R}$ and $K = [0,1]$, but now consider the (infinite) semigroup $\mathscr{S}$ generated by the pair of injective affine self-maps $\varphi(x) = (2x+1)/4$ and $\psi(x) = (x+1)/2$ of K. Since φ and ψ have no common fixed point, neither does $\mathscr{S}$. The exercise below shows what's wrong.

Exercise 13.2. Show that $\mathscr{S}$ as described above is not *uniformly* injective.

13.2 Extreme points of convex sets

For the proof of the Ryll–Nardzewski Theorem we'll make frequent use of the concepts of convex set, convex combination, and convex hull, as set out in Appendix C, Sect. C.1. Here's a crucial addition to this list.

Definition 13.3 (Extreme point)**.** For a convex subset C of a real vector space, an *extreme point* is a point of C that does not lie in the interior of the line segment joining two distinct points of C (i.e., a point that cannot be written as $tx + (1-t)y$, with $0 < t < 1$ and x, y distinct points of C).

Examples of extreme points. The endpoints of a closed interval of the real line, the vertices of a triangle in $\mathbb{R}^2$, or more generally a convex polygon in $\mathbb{R}^N$ (e.g., the standard simplex Π_N). Every point on the boundary of a closed ball in $\mathbb{R}^N$.

Non-examples. In a normed space: each point in the interior of a closed ball. For a convex polygon in $\mathbb{R}^N$: each point that is not a vertex (e.g., for Π_N, each point that is not one of the standard basis vectors for $\mathbb{R}^N$).

Exercise 13.3. Suppose C is a convex subset of a real vector space. Then $p \in C$ is an extreme point if and only if p cannot be represented nontrivially as a convex combination of other points of C.

A key step in our proof of the Ryll–Nardzewski Theorem will involve the following fundamental result about extreme points. If S is a subset of a topological vector space, we'll use the notation $\overline{\text{conv}}\, S$ for the closure of its convex hull.

Theorem 13.4 (A Krein–Milman theorem)**.** *Suppose K_0 is a nonempty compact subset of a locally convex topological vector space X and that $K = \overline{\text{conv}}\, K_0$ is also compact.*[5] *Then K_0 contains an extreme point of K.*

This result is a consequence of two famous theorems about nonempty compact subsets K of locally convex spaces. First there is *The* Krein–Milman Theorem, which asserts that not only does K *have* extreme points, it is in fact the *closed convex hull* of these extreme points. Next, the "Milman Inversion" of this theorem says that if K is the closed-convex hull of a compact subset K_0, then all of K's extreme points belong to K_0.[6] For our purposes we'll only need a special case of Theorem 13.4 (Theorem 13.6 below).

We begin with an even more special case of Theorem 13.4.

Lemma 13.5. *Suppose K_0 is a nonempty compact subset of an inner-product space. If $K := \overline{\text{conv}}\, K$ is compact then some point of K_0 is an extreme point of K.*

Proof. Let's denote the ambient inner-product space by X, its inner product by $\langle \cdot, \cdot \rangle$, and its norm by $\|\cdot\|$ (i.e., $\|x\|^2 = \langle x, x \rangle$ for each $x \in X$). Since K_0 is compact there is a smallest closed ball B in X that contains it, and so also contains the closure of its convex hull. Upon making an appropriate translation and dilation we may without loss of generality assume that B is the closed unit ball of X. The compactness of K_0 insures that it intersects ∂B in some vector v. This unit vector (or, for that matter, every unit vector in K_0) will turn out to be the desired extreme point for K. This is obvious from a picture; for an analytic proof suppose $v = tx + (1-t)y$ for some vectors $x, y \in K$ and for some $0 < t < 1$. Since $\|x\|$ and $\|y\|$ are both ≤ 1,

$$1 = \|v\|^2 = \langle v, v \rangle = \langle v, tx + (1-t)y \rangle = t\langle v, x \rangle + (1-t)\langle v, y \rangle$$

$$\leq t\|v\|\|x\| + (1-t)\|v\|\|y\| \leq 1,$$

where the next-to-last inequality follows from the Cauchy–Schwarz Inequality applied to both $\langle v, x \rangle$ and $\langle v, y \rangle$. Thus there is equality throughout, in particular

$$\|x\| = \|y\| = \langle v, x \rangle = \langle v, y \rangle = 1.$$

By the case of equality in the Cauchy–Schwarz inequality, this requires $x = \pm v$ and $y = \pm v$. They can't both be $-v$ lest $v = -v$, i.e., $v = 0$, contradicting the fact that v is a unit vector. On the other hand if one of them is v and the other is $-v$, then $v = \pm(2t - 1)v$ whereupon t is either 0 or 1, another contradiction. Thus v is an extreme point of K. $\square$

The version of Theorem 13.4 that we'll actually need is the following consequence of Lemma 13.5. Recall from Sect. 9.5 that if Y is a real vector space then

[5] Compactness of the closed convex hull of a compact set is automatic for Banach spaces (Proposition C.6, p. 195), but not so in general; it can fail even for non-closed subspaces of Hilbert space (Remark C.7, p. 195).

[6] For proofs of these theorems see, e.g., [103], Theorems 3.23 and 3.25, respectively, pp. 75–77.

the *weak-star topology* on the algebraic dual $Y^\sharp$ of Y is just the restriction to $Y^\sharp$ of the product topology of $\mathbb{R}^Y$. If Y is a *topological* vector space then its *dual space* Y^* is the collection of linear functionals on Y that are *continuous*. Note that Y^* is a linear subspace of $Y^\sharp$, and its weak-star topology is just the restriction of the weak-star topology of $Y^\sharp$, i.e., the topology of pointwise convergence on Y. We call Y the *predual* of Y^*.

Theorem 13.6 (A Krein–Milman Theorem for separable preduals)**.** *Suppose X is the dual of a separable Banach space, and $K_0 \subset X$ is nonempty and weak-star compact. If $\overline{\mathrm{conv}}\,K_0$ is weak-star compact, then it has an extreme point that lies in K_0.*

Proof. We are assuming that $X = Y^*$, where Y is a separable (real) Banach space. The closed unit ball Y_1 of Y has a countable dense subset $\{y_n\}_1^\infty$ (exercise). Each element of $f \in X$, being a continuous linear functional on Y, is bounded on Y_1. Consequently the formula

$$\langle f,g\rangle = \sum_{n=1}^{\infty} \frac{1}{2^n} f(y_n)g(y_n) \qquad (f,g \in X) \tag{13.2}$$

makes sense, and defines a bilinear form on X that is, in fact, an *inner product.* To see why, define $\|f\| := \sqrt{\langle f,f\rangle}$ for $f \in X$. To say that $\langle \cdot,\cdot\rangle$ is an inner product is to say that the seminorm $\|\cdot\|$ is a norm, i.e., that $\|f\| = 0$ only when $f = 0$. If $\|f\| = 0$ then by (13.2) we have $f(y_n) = 0$ for $n = 0,1,2,\ldots$, hence $f = 0$ on Y_1 by the continuity of f on Y and the density of $\{y_n\}_0^\infty$ in Y_1. Since f is a linear functional it must therefore vanish on all of Y.

CLAIM. *The norm topology η induced on X by this inner product coincides on weak-star compact sets with the weak-star topology ω.*

Once we've proved this Claim, the desired result on extreme points will follow from Lemma 13.5.

Proof of Claim. Let K be a weak-star compact subset of X. We need only show that the topology η is weaker than ω. Once this is done we'll know that the identity map j on K is continuous from ω to η, and so takes ω-closed subsets of K (which are ω-compact) to η-compact subsets of K (which are η-closed). Thus j is not just a continuous map from (K,ω) to (K,η), but also a *closed* one, and so (upon taking complements) an *open* one. Thus j is a homeomorphism, i.e., $\omega = \eta$.

To show that η is weaker than ω, note that by Proposition 9.10 (p. 111) we know that each vector $y \in Y_1$ induces an ω-continuous function $\hat{y}$: $K \to \mathbb{R}$ via the definition $\hat{y}(f) = f(y)$ $(f \in K)$. Since K is ω-compact, $\hat{y}$ is bounded thereon for each $y \in Y_1$. Turning things around: K is *pointwise bounded* on Y_1, hence by the Uniform Boundedness Principal[7] there exists a positive number M such that $|f| < M$ on Y_1 for every $f \in K$.

Now fix $\varepsilon > 0$ and a point $f_0 \in K$, and consider the relatively open subset U of K obtained by intersecting the η-ball of radius ε and center f_0 with K. Choose

[7] See, e.g., [102, Theorem 5.8, pp. 98–99], where it's called the "Banach-Steinhaus Principle."

a positive integer N for which $\sum_{n=N+1}^{\infty} 2^{-n} < \varepsilon^2/(8M^2)$. Let $F = \{y_0, y_1, \ldots, y_N\}$ be a finite subset of Y and suppose $f \in N(f_0, F, \varepsilon/\sqrt{2})$. Then, continuing with the notation $\|\cdot\|$ for the norm induced by the inner product (13.2):

$$\begin{aligned}
\|f - f_0\|^2 &= \sum_{n=1}^{\infty} \frac{1}{2^n}|f(y_n) - f_0(y_n)|^2 = \sum_{n=1}^{N} + \sum_{n=N+1}^{\infty} \frac{1}{2^n}|f(y_n) - f_0(y_n)|^2 \\
&< \frac{\varepsilon^2}{2} \sum_{n=1}^{N} \frac{1}{2^n} + (2M)^2 \sum_{n=N+1}^{\infty} \frac{1}{2^n} < \frac{\varepsilon^2}{2} + \varepsilon^2 \frac{4M^2}{8M^2} \\
&= \frac{\varepsilon^2}{2} + \frac{\varepsilon^2}{2} = \varepsilon^2.
\end{aligned}$$

Thus $N(f_0, F, \varepsilon/\sqrt{2}) \subset U$. We've shown that if $f_0 \in K$ then every η-neighborhood of f_0 contains an ω neighborhood of f_0, i.e., that the topology η induced on X by the inner product (13.2) is weaker than—and therefore equal to—the weak-star topology induced on X by its predual Y. □

The metrizability argument given above produces the following useful result:

Proposition 13.7 (Weak-star metrizability). *If X is the dual of a separable Banach space, then on each compact subset of X the weak-star topology is metrizable.*

In the next section we'll prove our "separable-predual" version of the Ryll–Nardzewski Theorem. The action will take place in the dual space of $C(G)$, where G is a compact, metrizable group, so we will need to know that $C(G)$ is separable. According to Proposition B.7 of Appendix B, this is true even if G is just a compact metric space.

13.3 Ryll–Nardzewski: separable predual version

In this section, X will be the (topological) dual of a separable (real) Banach space Y. Instead of considering X in its norm topology, however, we will endow it with the weak-star topology it gets from its predual Y. Thus by Proposition 13.7, every (weak-star) compact subset of X will be *metrizable.*

Theorem 13.8 ("Ryll–Nardzewski lite"). *Suppose X is the dual of a separable Banach space, K is a nonempty convex, weak-star compact subset of X, and $\mathscr{S}$ is a uniformly injective semigroup of continuous, affine self-maps of K. Then $\mathscr{S}$ has a common fixed point in K.*

Proof. The argument is best broken into several pieces.

Step I. It is enough to show:

(*) *Every finite subset of $\mathscr{S}$ has a common fixed point.*

For suppose we've established (*). If $A \in \mathscr{S}$ let F_A denote the fixed-point set of A:

$$F_A = \{x \in K : Ax = x\}.$$

We wish to show that $\bigcap\{F_A : A \in \mathscr{S}\}$ is nonempty. By the continuity of each map in $\mathscr{S}$ we know that each fixed-point set F_A is closed in K. Now $(*)$ is the assertion that the family of all these sets has the finite intersection property. Thus the compactness of K insures the entire family has nonvoid intersection.

Step II. Fix a finite subset $\mathscr{A} = \{A_1, A_2, \ldots A_n\}$ of $\mathscr{S}$. By Step I we'll be done if we can show that $\mathscr{A}$ has a common fixed point, so as noted above (Sect. 13.1, p. 164), we may as well assume that $\mathscr{S}$ is the semigroup generated by $\mathscr{A}$. Note that even though it is finitely generated, $\mathscr{S}$ need not be finite—if it were, we'd be done by Proposition 13.2. Nevertheless, as in the proof of that Proposition, we'll pin our hopes on the affine continuous self-map $A_0 = (A_1 + A_2 + \cdots + A_n)/n$ of K, for which the Markov–Kakutani Theorem once again guarantees a fixed point $x_0 \in K$. As in the case of finite $\mathscr{S}$, we'll show that x_0 is also a fixed point for each of the maps $A_1, \ldots, A_n$. Now, however, our argument needs to be more subtle.

Step III. Let $\mathscr{S}x_0 = \{Ax_0 : A \in \mathscr{S}\}$: the $\mathscr{S}$-orbit of x_0, and consider its closure K_0, a compact subset of K. Since $C = \overline{\text{conv}}\, K_0$ is a closed subset of K, it too is compact, so by Theorem 13.6 some point e of K_0 is an extreme point of C. Since K is metrizable (Proposition 13.7) there is a sequence (T_j) of maps in $\mathscr{S}$ such that $T_j x_0 \to e$. Since $A_0 x_0 = x_0$ we have

$$\begin{aligned} e &= \lim_j (T_j A_0) x_0 = \lim_j T_j \left(\frac{A_1 x_0 + A_2 x_0 + \cdots + A_n x_0}{n} \right) \\ &= \lim_j \left(\frac{(T_j A_1) x_0 + (T_j A_2) x_0 + \cdots + (T_j A_n) x_0}{n} \right). \end{aligned}$$

In the last line, which follows from the affine-ness of the maps T_j, we're looking at n sequences $((T_j A_k) x_0)_{j=1}^{\infty}$ for $k = 1, 2, \ldots, n$, each drawn from $\mathscr{S}x_0$. Thanks to the weak-star compactness and metrizability of K we can find a single subsequence (j_i) of indices such that the sequence $((T_{j_i} A_k) x_0)_{i=1}^{\infty}$ converges for each k to a vector $y_k \in K_0 = \overline{\mathscr{S}x_0}$. Thus the vector e, which we know belongs to $C = \overline{\text{conv}}\, K_0$, is actually the average of the vectors $y_1, y_2, \ldots, y_n \in K_0$ and so belongs to $\text{conv}\, K_0$. Since e is an extreme point of $\overline{\text{conv}}\, K_0$ the y_k's must all be equal to e.

Step IV. Recall that we're trying to show that $x_0 = A_k x_0$ for each $1 \le k \le n$. Since $A_0 x_0 = x_0$ it's enough to know (definition of A_0) that all the vectors $A_k x_0$ are the same. Choose two of them, say $A_\mu x_0$ and $A_\nu x_0$. We know from Step III that

$$0 = e - e = y_\mu - y_\nu = \lim_i [T_{j_i}(A_\mu x_0) - T_{j_i}(A_\nu x_0)],$$

so the zero vector belongs to the closure of the set $\{T(A_\mu x_0) - T(A_\nu x_0) : T \in \mathscr{S}\}$. This, plus the uniform injectivity of $\mathscr{S}$ (at last!), guarantees that $A_\mu x_0 = A_\nu x_0$. □

Ryll–Nardzewski at full strength (sketch of proof). To upgrade the proof just presented to one that establishes Theorem 13.1:

(a) In Step III: Instead of Theorem 13.6, the "separable predual" version of "A Krein–Milman Theorem," use the full-strength one, Theorem 13.4.

(b) In Step IV: Instead of sequences and subsequences, use "nets" and "subnets" (see the *Notes* below for the definition and discussion of these concepts). The proof just given then goes through *mutatis mutandis.* The reason for avoiding this generality is that, while nets provide a straightforward generalization of sequences, the same cannot be said for subnets vs. subsequences. In fact subnets of sequences need not be subsequences [123, Problem 11B, p. 77].

13.4 Application to Haar Measure

In the Ryll–Nardzewski Theorem we finally have a result that allows the duality method of Chaps. 9, 10, and 12 to establish the existence of Haar measure for *every* compact topological group—commutative or not. Recall the basics of this method: To each element γ of the group G we assign the left-translation operator L_γ defined on either $C(G)$ (the space of real-valued functions on G that are *continuous*) or $B(G)$ (the space of real-valued functions on G that are *bounded*) by $L_\gamma f(x) = f(\gamma x)$ $(x \in G)$. We implore a kind spirit to grant us a common fixed point for the collection of algebraic adjoints of these operators. The Riesz Representation Theorem then transforms this fixed point into Haar measure for G.

In Chap. 9 the group G was commutative and our translation adjoints lived on the algebraic dual of $C(G)$. In Chap. 10 we observed that the same argument worked as well in the algebraic dual of $B(G)$, where it produced an invariant mean which gave rise to an invariant finitely additive "probability measure" on *all* subsets of G. In Chap. 12, thanks to an enhanced Markov–Kakutani Theorem, the same method produced both invariant means and Haar measure for *solvable* compact groups. By invoking the Invariant Hahn–Banach Theorem (Theorem 12.5, p. 151) we could even produce an invariant mean whose associated finitely additive measure extended Haar measure from the Borel sets to all the subsets of G.

However our luck runs out if we try to extend the Markov–Kakutani Theorem further, in the hope of providing Haar measure for *all* compact groups. In Chap. 11 we saw that not every compact group has an invariant mean; the group $\mathrm{SO}(3)$ of rotation matrices, being paradoxical, furnishes just such an example. Thus, at least for $G = \mathrm{SO}(3)$, there's no kind spirit to provide an appropriate fixed point for translation adjoints on the algebraic dual of $B(G)$. However the story is different for the *topological* dual of $C(G)$, thanks to the Ryll–Nardzewski Theorem.

To apply that theorem we'll need to know that the continuity previously established for algebraic adjoints remains true of topological ones:

Lemma 13.9. *Suppose X is a Banach space and X^* its topological dual. If T is a continuous linear transformation on X then its topological adjoint T^* is weak-star continuous on X^*.*

Proof. The key here is that T^* is the restriction to X^* of the algebraic adjoint $T^\sharp$ acting on the algebraic dual $X^\sharp$. Proposition 9.18 tells us that $T^\sharp$ is weak-star continuous on $X^\sharp$. Now the weak-star topology on X^* is the restriction of the weak-star topology on $X^\sharp$, so T^* inherits the weak-star continuity of $T^\sharp$. □

Theorem 13.10. *Haar measure exists for every compact topological group.*

Outline of proof. Let $\mathscr{G}$ denote the group of left-translation operators L_γ for $\gamma \in G$, and let $\mathscr{G}^*$ be the corresponding group of adjoints, operating on $C(G)^*$. Let $\mathscr{K}$ denote the collection of positive linear functionals Λ on $C(G)$ with $\Lambda(1) = 1$.

By Exercise 9.14 (p. 114) we know that $|\Lambda(f)| \leq \|f\|$ for each $\Lambda \in \mathscr{K}$, hence $\mathscr{K}$ is a pointwise bounded subset of $C(G)^*$. By Theorem 9.12 (p. 111, our "infinite dimensional Heine–Borel theorem") $\mathscr{K}$ is therefore relatively compact in the product topology of $\mathbb{R}^{C(G)}$. Now $\mathscr{K}$ is the analogue for $C(G)$ of the set $\mathscr{M}$ of means on $B(G)$, and the proof that $\mathscr{M}$ is closed in $\mathbb{R}^{B(G)}$ works as well to show that $\mathscr{K}$ is closed in $\mathbb{R}^{C(G)}$, hence $\mathscr{K}$ is a compact subset of $\mathbb{R}^{C(G)}$. Since $\mathscr{K}$ is contained in $C(G)^*$, and since the weak-star topology of $C(G)^*$ is just the restriction to that space of the product topology of $\mathbb{R}^{C(G)}$, we see that *$\mathscr{K}$ is weak-star compact in $C(G)^*$.*

Clearly $\mathscr{K}$ is convex. Each of the operators in $\mathscr{G}^*$ is a linear self-map of $\mathscr{K}$ that, by Lemma 13.9, is weak-star continuous on $C(G)^*$. Thus if we can show that $\mathscr{G}^*$ is uniformly injective on $\mathscr{K}$, the Ryll–Nardzewski Theorem will provide a fixed point $\Lambda \in \mathscr{K}$ for $\mathscr{G}^*$. Just as in the commutative and solvable cases, the Riesz Representation Theorem will provide a regular Borel probability measure μ for G that represents Λ via integration, with the $\mathscr{G}^*$-invariance of Λ translating into left G-invariance for μ, i.e., μ will be Haar measure for G.

Proof for G metrizable. By Proposition B.7 we know that $C(G)$ is separable, hence Theorem 13.8, our "lite" version of the Ryll–Nardzewski Theorem, will apply to $C(G)^*$ (taken in its weak-star topology) once we've established that $\mathscr{G}^*$ is uniformly injective. For this we'll need to know that:

(†) *For each $\Lambda \in \mathscr{K}$ the map $\gamma \to L_\gamma^*\Lambda$ takes G continuously into $\mathscr{K}$ (with its weak-star topology).*

Proof. Let μ denote the regular Borel probability measure for G that—thanks to the Riesz Representation Theorem—represents Λ, i.e., $\Lambda(f) = \int_G f\,d\mu$ for $f \in C(G)$. Since both G and the weak-star topology on $\mathscr{K}$ are metrizable (the latter thanks to Propositions 13.7 and B.7) we may use sequences to establish continuity. Suppose (γ_n) is a sequence in G that converges to an element γ of G. Then for $f \in C(G)$ we have, thanks to the Dominated Convergence Theorem:

$$(L_{\gamma_n}^*\Lambda)(f) = \Lambda(L_{\gamma_n} f) = \int f(\gamma_n x)\,d\mu(x) \to \int f(\gamma x)\,d\mu(x) = (L_\gamma^*\Lambda)(f).$$

Thus $L_{\gamma_n}^*\Lambda \to L_\gamma^*\Lambda$ in the weak-star topology (cf. Exercise 9.10, p.110), which establishes the desired continuity of the map $\gamma \to L_\gamma^*$. □

To show is that $\mathscr{G}^*$ is uniformly injective on $\mathscr{K}$, fix Φ and Λ in $C(G)^*$ and suppose the zero-functional belongs to the weak-star closure of $\Delta = \{L_\gamma^*\Phi - L_\gamma^*\Lambda : \gamma \in G\}$. Our goal is to prove that $\Phi = \Lambda$. Note that the set Δ is pointwise bounded on $C(G)$, so its closure is weak-star compact, and therefore metrizable. This, along with the metrizability of G allows the following rephrasing of hypothesis on Δ: *There exists a sequence (γ_n) of elements of G such that $L_{\gamma_n}^*\Phi - L_{\gamma_n}^*\Lambda \to 0$ weak-star in $C(G)^*$.* Since G is compact we may, upon replacing our original sequence of group elements by an appropriate subsequence, assume that (γ_n) converges to an element $\gamma \in G$. By the continuity established in (†) above we have

$$0 = \lim_n [L_{\gamma_n}^*\Phi - L_{\gamma_n}^*\Lambda] = L_\gamma^*\Phi - L_\gamma^*\Lambda = \Phi \circ L_\gamma - \Lambda \circ L_\gamma = (\Phi - \Lambda) \circ L_\gamma.$$

Since L_γ is an isomorphism of $C(G)$ onto itself this implies $\Phi = \Lambda$, as desired.

Thus the hypotheses of our "lite" version of the Ryll–Nardzewski Theorem (Theorem 13.8) are satisfied with $X = C(G)^*$, $K = \mathscr{K}$, and $\mathscr{S} = \mathscr{G}^*$, so $\mathscr{G}^*$ has a fixed point in $\mathscr{K}$; as noted above, this provides Haar measure for G. □

Sketch of proof for arbitrary compact G. In this setting the weak-star topology on $C(G)^*$ is no longer metrizable on every compact set, so we can't use sequential arguments. This means we must modify the continuity proof for $\gamma \to L_\gamma^*$ so as to avoid the Dominated Convergence Theorem. Instead the idea is to first prove that for each $f \in C(G)$ the map $\gamma \to L_\gamma f$ is continuous from G to $C(G)$ in its norm topology. This follows from the fact each continuous function on G exhibits a form of uniform continuity that generalizes the one familiar to us from metric-space theory.[8] Once we've established the desired continuity of the map $\gamma \to L_\gamma f$ it's an easy matter to show that the map $\gamma \to L_\gamma^*\Lambda$ is continuous for each $\Lambda \in C(G)^*$. The rest of the argument then goes through almost word-for-word, with nets and subnets replacing sequences and subsequences.

Now that we know Haar measure exists (uniquely and bi-invariantly) on every compact group, it's time to investigate an important example.

13.5 Haar Measure on SO(3)

SO(3) is the collection of 3×3 matrices whose determinant is 1 and whose columns form an orthonormal subset of $\mathbb{R}^3$. For every real square matrix A, such column orthonormality expresses itself in the matrix equation $AA^t = I$, where A^t denotes the transpose of A, and I is the identity matrix of the size of A. This, along with the multiplicative property of determinants, makes it easy to show that SO(3) is a group under matrix multiplication; in Appendix D it's shown that the elements of this group are precisely the matrices (with respect to the standard basis of $\mathbb{R}^3$) of rota-

[8] See, e.g., [103], proof of Theorem 5.13, pp. 129–130.

tions of $\mathbb{R}^3$ about the origin. For topological purposes we'll regard $\mathrm{SO}(3)$ as a subset of the sphere of radius $\sqrt{3}$ in $\mathbb{R}^9$. It's easy to check that the Euclidean topology of $\mathbb{R}^9$ makes $\mathrm{SO}(3)$ into a compact group (exercise), which therefore possesses Haar measure. What is this measure? How does one integrate respect to it?

What is Haar Measure on SO(3)? Since we can regard $\mathrm{SO}(3)$ as the group of rotations of the unit sphere S^2 of $\mathbb{R}^3$, one might suspect that its Haar measure should somehow involve surface area measure on that sphere. A natural way of connecting group with sphere is to define the map $\varphi\colon \mathrm{SO}(3) \to S^2$ which takes a matrix $x \in \mathrm{SO}(3)$ to its last column. Thus $\varphi(x) = xe_3$, where we regard $\mathbb{R}^3$ as a space of column vectors, and $e_3 = [0,0,1]^t$ is the unit vector in $\mathbb{R}^3$ "along the z-axis."

Since matrix entries are continuous functions of their matrices, the map φ is continuous. It is surjective (each unit vector can be the third column of a matrix in $\mathrm{SO}(3)$), but not one-to-one (it's constant on subsets of $\mathrm{SO}(3)$ whose elements share the same third column).

More precisely, let K denote the subgroup of matrices in $\mathrm{SO}(3)$ that fix the vector e_3 (i.e., which have third column equal to e_3). Then the coset modulo K of a matrix $x \in \mathrm{SO}(3)$ is $xK = \{xk : k \in K\}$, namely all matrices in $\mathrm{SO}(3)$ with third column the same as that of x. If x and y in $\mathrm{SO}(3)$ have different third columns (i.e., belong to different cosets mod K), then $\varphi(x) \neq \varphi(y)$. Thus φ is a one-to-one mapping of cosets mod K onto S^2.[9]

Now suppose $f \in C\big(\mathrm{SO}(3)\big)$. The subgroup K, being compact, has its own Haar measure which we'll denote by dk. Define f_K on $C\big(\mathrm{SO}(3)\big)$ by:

$$f_K(x) := \int_K f(xk)dk \qquad \big(x \in \mathrm{SO}(3)\big).$$

Clearly f_K is continuous; by the invariance of dk it is constant on cosets of $\mathrm{SO}(3)$ modulo K.

To make the definition of f_K more concrete, observe that each element of K has the form

$$k(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad (0 \le \theta < 2\pi),$$

so the map that takes $k(\theta)$ to its upper left-hand 2×2 submatrix, and then to the unimodular complex number $\cos\theta + i\sin\theta$, establishes a homeomorphic isomorphism between K and the unit circle, now viewed as the group of rotations of $\mathbb{R}^2$. This allows Haar measure on K to be concretely represented by the Haar measure of the circle group, i.e., normalized Lebesgue arc-length measure:

$$\int_K g(k)\,dk = \frac{1}{2\pi}\int_0^{2\pi} g(k(\theta))\,d\theta \qquad \big(g \in C(K)\big).$$

[9] As such, φ can be regarded as a map taking $\mathrm{SO}(3)$ onto the quotient space $\mathrm{SO}(3)/K$, but right now we'll avoid the notion of "quotient space."

Thus for $f \in C\big(\mathrm{SO}(3)\big)$,

$$f_K(x) = \frac{1}{2\pi}\int_0^{2\pi} f\big(xk(\theta)\big)\,d\theta \qquad \big(x \in \mathrm{SO}(3)\big). \tag{13.3}$$

Exercise 13.4. Define $f\colon \mathrm{SO}(3) \to [-1,1]$ by $f(x) = (x_{2,2})^2$. Show that

$$f_K(x) = \frac{1-(x_{2,3})^2}{2}$$

for each $x \in \mathrm{SO}(3)$.

The function f_K, being constant on cosets mod K, may be viewed via the map φ as a function on S^2. More precisely, let

$$\hat{f}(p) = f_K\big(\varphi^{-1}(p)\big) \qquad (p \in S^2).$$

Theorem 13.11. *Haar measure dx on* $\mathrm{SO}(3)$ *is given by*

$$\int_{\mathrm{SO}(3)} f(x)\,dx = \int_{S^2} \hat{f}(p)\,d\sigma(p) \qquad \big(f \in C\big(\mathrm{SO}(3)\big)\big),$$

where σ denotes surface area measure on S^2, normalized to have unit mass.

Proof. Define the linear functional Λ on $C\big(\mathrm{SO}(3)\big)$ by

$$\Lambda(f) = \int_{S^2} \hat{f}(p)\,d\sigma(p) \qquad \big(f \in C\big(\mathrm{SO}(3)\big)\big).$$

Then Λ is a positive linear functional on $C\big(\mathrm{SO}(3)\big)$, so the Riesz Representation Theorem provides a regular Borel probability measure μ for $\mathrm{SO}(3)$ such that $\Lambda(f) = \int f\,d\mu$ for all $f \in C\big(\mathrm{SO}(3)\big)$. One checks easily that if $f \equiv 1$ on $\mathrm{SO}(3)$ then $\Lambda(f) = 1$, i.e., that $\mu\big(\mathrm{SO}(3)\big) = 1$; thus μ is a probability measure.

To show that μ is Haar measure on $\mathrm{SO}(3)$ we need only check is that it is left-invariant, i.e., that if L_y is the "left-translation" operator on $C\big(\mathrm{SO}(3)\big)$:

$$(L_y f)(x) = f(yx) \qquad \big(f \in C\big(\mathrm{SO}(3)\big),\ x,y \in \mathrm{SO}(3)\big),$$

then $\Lambda(L_y f) = \Lambda(f)$ for $f \in C\big(\mathrm{SO}(3)\big)$ and $y \in \mathrm{SO}(3)$.

The proof of this hinges on the identity

$$\widehat{(L_y f)} = L_y(\hat{f}) \qquad \text{for each } f \in C\big(\mathrm{SO}(3)\big), \tag{13.4}$$

where on the right-hand side we have L_y operating in the obvious way on $C(S^2)$, namely: $L_y g(p) = g(yp)$ for $p \in S^2$ and $g \in C(S^2)$. Granting this: for $f \in C\big(\mathrm{SO}(3)\big)$ and $y \in \mathrm{SO}(3)$:

$$\Lambda(L_y f) = \int_{S^2} \widehat{L_y f}\,d\sigma = \int_{S^2} L_y(\hat{f})\,d\sigma = \int_{S^2} \hat{f}\,d\sigma = \Lambda(f),$$

where the second equality uses (13.4), and the third one follows from the rotation-invariance of surface area measure on S^2.

The proof of (13.4) involves nothing more than chasing definitions. Fix $p \in S^2$ and choose $x \in \mathrm{SO}(3)$ with $\varphi(x) = p$ (so that p is the third column of the matrix x). Fix the "translator" $y \in \mathrm{SO}(3)$. Note that

$$\varphi(yx) := yxe_3 = y\varphi(x) = yp \tag{13.5}$$

so for $f \in C\big(\mathrm{SO}(3)\big)$,

$$\begin{aligned}
\widehat{(L_y f)}(p) = (L_y f)_K\big(\varphi^{-1}(p)\big) &= \int_K (L_y f)(xk)\,dk \\
&= \int_K f(yxk)\,dk = f_K(yx) \\
&= f_K(\varphi^{-1}(yp)) \qquad \text{(by (13.5))} \\
&= \hat{f}(yp) = (L_y \hat{f})(p)
\end{aligned}$$

which completes the proof of the theorem. □

It's tempting to think of Theorem 13.11 as somehow expressing Haar measure on $\mathrm{SO}(3)$ as the product of normalized surface area on the sphere S^2 and Haar measure on the subgroup K. Not so! One must instead regard Theorem 13.11 as "disintegrating" Haar measure on $\mathrm{SO}(3)$ into a family of translates of dk—one for each coset mod K of $\mathrm{SO}(3)$—which are "glued together" by the surface area measure $d\sigma$.

More precisely, denote left-multiplication by x on $\mathrm{SO}(3)$ by λ_x (i.e., $\lambda_x(y) = xy$ for $y \in \mathrm{SO}(3)$), and Haar measure on K by ν. Then for $x \in \mathrm{SO}(3)$ can use the change-of-variable formula of measure theory to rewrite the definition of $f_K(x)$ as:

$$f_K(x) = \int_K f\big(\lambda_x(k)\big)\,d\nu(k) = \int_{xK} f\,d(\nu\lambda_x^{-1}) \qquad \big(f \in C\big(\mathrm{SO}(3)\big)\big).$$

We have $xK = \varphi^{-1}(p)$, where p is the third column of the matrix x. Since $f_K(x)$ is constant on $\varphi^{-1}(p)$, the formula above shows that the probability measure $\nu\lambda_x^{-1}$ depends only on the point $p \in S^2$. Upon writing ν_p for this measure we can rewrite the conclusion of Theorem 13.11 as:

$$\int_{\mathrm{SO}(3)} f(x)\,dx = \int_{S^2} \left(\int_{\varphi^{-1}(p)} f\,d\nu_p\right) d\sigma(p) \qquad \big(f \in C\big(\mathrm{SO}(3)\big)\big), \tag{13.6}$$

which exhibits how Haar measure on $\mathrm{SO}(3)$ "disintegrates" into the measures ν_p.

Exercise 13.5. Express the familiar formula from Calculus by which one integrates a continuous real-valued function over the plane triangle $\Delta = \{0 \le y \le x, 0 \le x \le 1\}$ as a similar "disintegration" of Lebesgue area measure on Δ into a family of one dimensional measures on vertical (or, if you wish, the horizontal) cross-sections of that triangle.

Exercise 13.6. Identify Haar measure on O(3), the group of *all* orthogonal 3×3 matrices.

13.6 Computation of some Haar integrals over SO(3)

Let's use the usual subscript notion $x_{i,j}$ to denote the entry in the i-th row and j-th column of a matrix x. Suppose g is a real-valued continuous function on the closed real interval $[-1,1]$. What is $\int_{\mathrm{SO}(3)} g(x_{i,j})\,dx$? For $i = j = 3$ the answer is easy to find, since in this case the function $g(x) = g(x_{3,3})$ is already constant on cosets of $\mathrm{SO}(3)$ mod K, hence in our characterization of Haar measure on $\mathrm{SO}(3)$, $g = g_K$, and therefore

$$\int_{\mathrm{SO}(3)} g(x_{3,3})\,dx = \int_{S^2} g(p_3)\,d\sigma(p). \tag{13.7}$$

Let (θ,φ) be the usual spherical coordinates of a point of $p \in S^2$, i.e., φ is the angle from the z-axis to the line from the origin to p, and θ is the angle from the x-axis to that line. Thus $p = [\cos\theta\sin\varphi, \sin\theta\sin\varphi, \cos\varphi]^t$. In particular, $p_3 = \cos\varphi$, and we know from multivariable calculus that (normalized) area measure on S^2 is given by $d\sigma(p) = \frac{1}{4\pi}\sin\varphi\,d\varphi\,d\theta$. Thus the right-hand side of (13.7) is

$$\frac{1}{4\pi}\int_{\theta=0}^{2\pi}\left(\int_{\varphi=0}^{\pi} g(\cos\varphi)\sin\varphi\,d\varphi\right)d\theta\,,$$

so upon setting $t = -\cos\varphi$ in the inner integral we obtain

$$\int_{\mathrm{SO}(3)} g(x_{3,3})\,dx = \frac{1}{2}\int_{-1}^{1} g(t)\,dt\,. \tag{13.8}$$

The same reasoning could be used to integrate $g(x_{i,3})$ for $i = 1,2$, but there's no reason to do so; the bi-invariance of Haar measure reduces all such integrals to the one we just worked out, yielding

Proposition 13.12. *Suppose* $g \in C([-1,1])$ *and* $1 \le i,j \le 3$. *Then*

$$\int_{\mathrm{SO}(3)} g(x_{i,j})\,dx = \frac{1}{2}\int_{-1}^{1} g(t)\,dt\,.$$

Proof. One can find matrices $a,b \in \mathrm{SO}(3)$ such that $x_{i,j} = (axb)_{3,3}$ (exercise) whereupon the bi-invariance of Haar measure on $\mathrm{SO}(3)$ yields

$$\int_{\mathrm{SO}(3)} g(x_{i,j})\,dx = \int_{\mathrm{SO}(3)} g\big((axb)_{3,3}\big)\,dx = \int_{\mathrm{SO}(3)} g(x_{3,3})\,dx\,.$$

This, along with (13.8), above gives the promised result. □

Corollary 13.13. $\int_{\mathrm{SO}(3)} x_{i,j}\,dx = 0$ *for all* (i,j) *with* $1 \le i,j \le 3$.

Exercise 13.7. Show that the "normalized" matrix entries $\{x_{i,j}/\sqrt{3} : 1 \le i, j \le 3\}$ form an orthonormal set in $L^2(\mathrm{SO}(3))$ (with respect to Haar measure).

Suggestion. To prove orthogonality, use the bi-invariance of Haar measure to reduce the problem to showing that $x_{i,j} \perp x_{3,3}$ whenever $(i,j) \neq (3,3)$. For this you'll need to show that if $f(x) = x_{i,j}$ then $f_K(x) = 0$ whenever $j \neq 3$, and $= x_{i,3}$ otherwise.

Characters again. Recall from Sect. 9.7 the notion of *character* for a topological group: a continuous homomorphism of that group into the circle group $\mathbb{T}$. We saw in that section that characters form the basis of an extension to compact *abelian* groups of Fourier analysis on the circle. The exercise below shows that the situation is much different for non-commutative groups.

Exercise 13.8 (The character group of SO(3) is trivial). This exercise requires only the fact that matrices in SO(3) are in one-to-one correspondence with rotations of $\mathbb{R}^3$ about the origin, that rotations preserve lengths of vectors and angles between vectors, and that each such rotation is uniquely determined by its axis (a line through the origin, each point of which is fixed) and its angle of rotation about that axis. For full details see Appendix D.

Let $R_u(\theta)$ denote the rotation having axis in the direction of the unit vector u and rotation angle $\theta \in [-\pi, \pi)$, where the sign of the angle is determined by the "right-hand rule." If u is the unit vector along the x-axis, we'll write $R_x(\theta)$ instead of $R_u(\theta)$, and similarly for $R_z(\theta)$.

(a) Suppose u, v is a pair of unit vectors in $\mathbb{R}^3$, and $M \in \mathrm{SO}(3)$ maps u on to v. Then for each angle θ we have the (unitary) similarity $R_v(\theta) = MR_u(\theta)M^{-1}$ (see Appendix D, p. 205 for a more detailed version of this). Conclude that for each character γ on SO(3), the value at each matrix in SO(3) depends only on the angle of rotation and not on the axis.

(b) Let $M(\theta) = R_z(\theta)R_x(-\theta)$. Show that if γ is a character of SO(3) then $\gamma(M(\theta)) = 1$ for every $\theta \in [-\pi, \pi)$. Thus one need only prove that every $\alpha \in [-\pi, \pi)$ is the angle of rotation of some $M(\theta)$ or its inverse.

(c) Prove that $\cos\theta = \big(\mathrm{trace}\,(R_u(\theta)) - 1\big)/2$. Use this to show that if $f(\theta)$ is the cosine of the angle of rotation of $M(\theta)$ then

$$f(\theta) = -1/4 + \cos(t) + \cos(2t)/4.$$

Show that f maps the interval $(-\pi, \pi]$ onto $[-1, 1]$. Conclude that if $\beta \in (-\pi, \pi]$ then there exists $\theta \in (-\pi, \pi]$ such that either $M(\theta)$ or its inverse is a rotation through angle β.Thus SO(3) has only the trivial character $\gamma \equiv 1$.

13.7 Kakutani's Equicontinuity Theorem

The Ryll–Nardzewski Theorem generalizes:

Corollary 13.14 (Kakutani)**.** *Suppose K is a nonvoid, compact, convex subset of a normed linear space, and $\mathscr{G}$ is an* equicontinuous *group of affine self-maps of K. Then $\mathscr{G}$ has a fixed point in K.*

Proof. By Ryll–Nardzewski's Theorem it's enough to prove that the group $\mathscr{G}$ is uniformly injective. To this end suppose x and y are vectors in K with $x \neq y$. Let $\varepsilon = \|x - y\|$. By equicontinuity there exists $\delta > 0$ such that if v and w are vectors in K with $\|v - w\| < \delta$ then $\|A(v) - A(w)\| < \varepsilon$ for every $A \in \mathscr{G}$. Now fix $A \in \mathscr{G}$, so A^{-1} also belongs to $\mathscr{G}$ and $\varepsilon = \|A^{-1}A(x) - A^{-1}A(y)\|$, hence $\|A(x) - A(y)\|$ must be $\geq \delta$. Thus the zero vector does not belong to the closure of $\{A(x) - A(y) \colon A \in \mathscr{G}\}$, which establishes the uniform injectivity of $\mathscr{G}$. □

This proof, with the notion of "equicontinuity" suitably interpreted, can be made to work as well in every locally convex topological vector space; see, for example, [103, Theorem 5.11, pp. 127–128]. Kakutani's theorem can be used as the first step of a proof (much different from the one given above) of the existence of Haar measure for every compact group. See, e.g., [103, Theorems 5.13–5.14, pp. 129–132].

Notes

The "real" Ryll–Nardzewski Theorem. Our version of Ryll–Nardzewski's theorem (Theorem 13.1) is due to Hahn [45]; it's a special case of what Ryll–Nardzewski actually proved. The "real" result, proved in [105], assumes compactness of the convex set K and continuity for the affine semigroup of maps $\mathscr{S}$ for the *weak topology* induced on the locally convex space X by its dual space. The notion of "uniform injectivity" however still refers to the original topology of X. Shorter proofs were subsequently given by Namioka and Asplund [82], and later by Dugundji and Granas [35].

Nets. A *sequence* $(x_n)_1^\infty$ from a set X is just a function x from the set $\mathbb{N}$ of natural numbers to X, with x_n denoting the value $x(n)$. More generally, suppose D is a set on which there is a relation $\prec$ that is both reflexive and transitive, and for which every pair of elements in A has an upper bound. The pair $(A, \prec)$ is called a *directed set*, and a function $x \colon A \to X$ is called a *net* from X, often abbreviated $(x_\delta)_{\delta \in D}$. If X is a topological space then to say such a net *converges* to an element $x_0 \in X$ means that for every neighborhood U of x_0 there exists $\delta_0 \in D$ such that $\delta_0 \prec \delta \implies x_\delta \in U$. With this definition, the sequential arguments that establish the properties of closure and continuity for metric spaces can be carried over directly to general topological spaces simply by replacing sequences with nets. The subsequential characterization of compactness for metric spaces even has an analogue for nets in general topological spaces, but for this to happen the proper definition of "subnet" must be a lot more subtle than that of "subsequence." For the details see, e.g., [123, Chap. 4].

Haar Measure is named for the Hungarian mathematician Alfred Haar (1885–1933) whose landmark paper [42, 1933] proved its existence for metrizable locally compact groups. Subsequently Banach [8, 1937] modified Haar's argument to provide measures invariant for the action of compact transformation groups acting continuously on compact metric spaces. The existence of an invariant measure for each locally compact (Hausdorff) topological group was proved in 1940 by André Weil. Predating all of this, in 1897 Adolph Hurwitz defined the notion of invariant integral for SO(n), essentially identifying Haar measure for that group. For this, and further historical background and references, see Hawkins' exposition [49, 1999] (especially p. 185 for Hurwitz's result, and pp. 194–196 for the rest). Diestel and Spalsbury in [29, 2014] provide a recent and accessible account of Haar measure, its history, and many of its applications including a nice introduction to its role in harmonic analysis on compact groups as well as some recent applications to Banach space theory.

Disintegration of Haar Measure with respect to a subgroup. The argument that proved Theorem 13.11 goes through almost *verbatim* to prove the same result with SO(3) replaced by a compact group G, K by a closed subgroup, S^2 by G/K, and σ by $\mu\pi^{-1}$, where $\pi\colon G \to G/K$ is the "quotient map" $x \to xK$ $(x \in G)$. The point is that the quotient space G/K has a natural topology that renders it compact and Hausdorff (namely: the strongest topology that makes π continuous); in the case of SO(3) this is just the topology induced on G/K by its identification with S^2. One also needs to note that the natural action of G on G/K ($xK \to gxK$ for $g, x \in K$) is continuous in this topology. With these substitutions Theorem 13.11 remains true, and signals a disintegration of μ with respect to the family of translates ν_p of ν to the cosets p that make up G/K.

Exercise 13.8. The argument outlined for this exercise expands on that of [36, Sect. 4.8.4, p. 232].