Chapter 4
Probability Measures on Compact Lie Groups

Abstract We introduce the space of Borel probability measures (and the important subspaces of central and symmetric measures) on a group, and topologise this space with the topology of weak convergence. A key tool for studying such measures is the (non-commutative) Fourier transform, which we extend from its action on functions that we described in Chap. 2. We discuss Lo-Ng positivity as a possible replacement for Bochner's theorem in this context. The theorems of Raikov-Williamson and Raikov are presented that give necessary and sufficient conditions for absolute continuity with respect to Haar measure. We then use the Fourier transform to find conditions for square-integrable densities, and the Sugiura space techniques of Chap. 3 to investigate smoothness of densities. Next we turn our attention to classifying idempotent measures and present the Kawada-Itô equidistribution theorem for the convergence of convolution powers of a measure to the uniform distribution. We introduce and establish key properties of convolution operators, including the notion of associated (sub/super-)harmonic functions. Finally we study some properties of recurrent measures on groups.

4.1 Classes of Probability Measures and Convolution

Let $\mathcal{P}(G)$ be the set of all Borel probability measures defined on an arbitrary Lie group G. As discussed in Appendix A.5, every $\mu \in \mathcal{P}(G)$ is both regular and Radon.[1] We equip $\mathcal{P}(G)$ with the topology of *weak convergence*, so if $(\mu_n, n \in \mathbb{N})$ is a sequence of measures in $\mathcal{P}(G)$ and $\mu \in \mathcal{P}(G)$, we say that the sequence converges to μ weakly as $n \to \infty$ if $\lim_{n\to\infty} \int_G f(x)\mu_n(dx) = \int_G f(x)\mu(dx)$ for all $f \in C_b(G, \mathbb{R})$. In this case we sometimes write $\mu_n \underset{w}{\to} \mu$ as $n \to \infty$. In the Chap. 5

[1] But if we drop the condition that G be a Lie group, we should work instead with *regular* Borel probability measures on the topological group G.

D. Applebaum, *Probability on Compact Lie Groups*, Probability Theory and Stochastic Modelling 70, DOI: 10.1007/978-3-319-07842-7_4,

we will also need *vague convergence* of probability measures and this is defined in exactly the same manner as weak convergence, except that the part of $C_b(G, \mathbb{R})$ is played by $C_c(G, \mathbb{R})$.

If $\mu \in \mathcal{P}(G)$, its *reversed measure* is $\widetilde{\mu} \in \mathcal{P}(G)$ where $\widetilde{\mu}(A) := \mu(A^{-1})$ for all $A \in \mathcal{P}(G)$. We say that $\mu \in \mathcal{P}(G)$ is *symmetric* if $\mu = \widetilde{\mu}$ and *central* (or *conjugate invariant*) if $\mu(gAg^{-1}) = \mu(A)$ for all $g \in G$ and all $A \in \mathcal{B}(G)$. We write $\mathcal{P}_s(G)$ and $\mathcal{P}_c(G)$ to denote the spaces of symmetric and central Borel probability measures defined on G (respectively), and we define $\mathcal{P}_{sc}(G) = \mathcal{P}_c(G) \cap \mathcal{P}_s(G)$.

If we are given a (left or right) Haar measure on G (which is always, as usual, assumed to be normalised when G is compact), we define $\mathcal{P}_{ac}(G)$ to be the corresponding subset of $\mathcal{P}(G)$ comprising absolutely continuous measures, so $\mu \in \mathcal{P}_{ac}(G)$ if there exists $f_\mu \in L^1(G)$ so that $\mu(A) = \int_A f_\mu(g)dg$ for all $A \in \mathcal{B}(G)$. The Radon-Nikodym derivative f_μ is called the *density* of the measure μ (with respect to the given Haar measure). If G is compact and $\mu \in \mathcal{P}_{ac}$, then $\mu \in \mathcal{P}_s$ if and only if $f_\mu(g) = f_\mu(g^{-1})$ for almost all $g \in G$, and $\mu \in \mathcal{P}_c$ if and only if for all $h \in G$, $f_\mu(hgh^{-1}) = f_\mu(g)$ for almost all $g \in G$.

To see that $\mathcal{P}(G) \neq \emptyset$, consider the *Dirac mass* δ_g at the point $g \in G$ which is defined for each $A \in \mathcal{B}(G)$ by

$$\delta_g(A) = \begin{cases} 1 & \text{if } g \in A \\ 0 & \text{if } g \notin A \end{cases}.$$

Clearly $\delta_g \in \mathcal{P}(G)$ and $\widetilde{\delta_g} = \delta_{g^{-1}}$. We may also form measures in $\mathcal{P}(G)$ by taking convex combinations of distinct Dirac masses. We will consider many more interesting examples as this and the subsequent chapter unfold.

Let $\mu_1, \mu_2 \in \mathcal{P}(G)$. Using the Riesz representation theorem we may assert the existence in $\mathcal{P}(G)$ of the left and right *convolution* products $\mu_1 *_L \mu_2$ and $\mu_1 *_R \mu_2$, which are defined (respectively) for all $f \in C_c(G)$ by

$$\int_G f(g)(\mu_1 *_L \mu_2)(dg) = \int_G \int_G f(gh)\mu_1(dg)\mu_2(dh),$$

$$\int_G f(g)(\mu_1 *_R \mu_2)(dg) = \int_G \int_G f(hg)\mu_1(dg)\mu_2(dh).$$

It is easily verified that $\mu_1 *_R \mu_2 = \mu_2 *_L \mu_1$. From now on we will only deal with left convolution, and we will write $\mu_1 * \mu_2 := \mu_1 *_L \mu_2$. It can be shown (see e.g. Stromberg [197]) that for all $B \in \mathcal{B}(G)$

$$\begin{aligned}(\mu_1 * \mu_2)(B) &= \int_G \int_G 1_B(gh)\mu_1(dg)\mu_2(dh) \\ &= \int_G \mu_1(Bh^{-1})\mu_2(dh) = \int_G \mu_2(g^{-1}B)\mu_1(dg).\end{aligned} \tag{4.1.1}$$

Convolution is associative, i.e. if $\mu_1, \mu_2, \mu_3 \in \mathcal{P}(G)$, then $(\mu_1 * \mu_2) * \mu_3 = \mu_1 * (\mu_2 * \mu_3)$, and so $(\mathcal{P}(G), *)$ is a semigroup . But note that if G is not abelian, we cannot expect commutativity to hold. Indeed, you can easily check that for $g, h \in G$, $\delta_g * \delta_h = \delta_{gh}$, and so $\delta_g * \delta_h = \delta_h * \delta_g$ if and only if $gh = hg$. In the general case $(\mathcal{P}(G), *)$ is a monoid (i.e. a semigroup with an identity element), since $\mu * \delta_e = \delta_e * \mu$ for all $\mu \in \mathcal{P}(G)$.

If $\mu \in \mathcal{P}_{ac}(G)$ and $\nu \in \mathcal{P}(G)$, we write $f_\mu * \nu := \mu * \nu$ and $\nu * f_\mu := \nu * \mu$. By using Fubini's theorem we easily verify that if we employ a right-invariant Haar measure, then $f_\mu * \nu \in \mathcal{P}_{ac}(G)$ with density $\int_G f_\mu(gh^{-1})\nu(dh)$ and if we choose a left-invariant Haar measure, then $\nu * f_\mu \in \mathcal{P}_{ac}(G)$ with density $\int_G f_\mu(g^{-1}h)\nu(dh)$.

The operation $\widetilde{\cdot}$ acts as an involution on $(\mathcal{P}(G), *)$. Indeed, we have $\widetilde{\widetilde{\mu}} = \mu$ for all $\mu \in \mathcal{P}(G)$, $\widetilde{\mu_1 * \mu_2} = \widetilde{\mu_2} * \widetilde{\mu_1}$ for all $\mu_1, \mu_2 \in \mathcal{P}(G)$ and $\widetilde{\delta_e} = \delta_e$.

The *support* of $\mu \in \mathcal{P}(G)$, which we denote by $\text{supp}(\mu)$, is the set of all $g \in G$ for which every Borel neighbourhood of g has strictly positive μ-measure. It is clear that $\text{supp}(\mu)$ is a closed subset of G. It is shown in Wendel [217] (pp. 925–926) that if μ_1, μ_2 are regular probability measures on G, then

$$\text{supp}(\mu_1 * \mu_2) = \text{supp}(\mu_1)\text{supp}(\mu_2), \tag{4.1.2}$$

where if $A, B \in \mathcal{B}(G)$, $AB := \{gh, g \in A, h \in B\}$ (and for later usage $A^2 := AA$).

Although we won't use it in the sequel, the next result may be of interest.

Proposition 4.1.1 *If G is a compact group, then the space $\mathcal{P}(G)$, equipped with the weak topology, is compact.*

Proof By identifying each $\mu \in \mathcal{P}(G)$ with the linear functional I_μ on $C(G, \mathbb{R})$ defined by $I_\mu(f) = \int_G f(g)\mu(dg)$ for $f \in C(G, \mathbb{R})$, we embed $\mathcal{P}(G)$ into the topological dual space $C(G, \mathbb{R})^*$, and recognise that the weak topology on $\mathcal{P}(G)$ is in fact the restriction of the weak-$*$ topology on $C(G, \mathbb{R})^*$. By the Banach-Alaoglu theorem, the unit ball in $C(G, \mathbb{R})^*$ is weak-$*$ compact. However, $\mathcal{P}(G)$ is easily verified to be a closed subset of this ball, and the result follows. □

Note that the mapping $g \to \delta_g$ is a continuous embedding of G into a closed subspace of $\mathcal{P}(G)$.

We recall that a family of Borel probability measures $(\mu_\alpha \in \mathcal{I})$ defined on some locally compact space X (where $\mathcal{I}$ is some index set) is *tight* if given any $\epsilon > 0$ there exists a compact set K_ϵ such that $\mu_\alpha(K_\epsilon) > 1 - \epsilon$ for all $\alpha \in \mathcal{I}$. If X is itself compact, then it is clear that any family of probability measures is tight (just take $K_\epsilon = X$ for all ϵ). So on a compact group G, by Prohorov's theorem, (see e.g. Heyer [95] Theorem 1.1.11, p. 26), any family of Borel probability measures $(\mu_\alpha \in \mathcal{I})$ contains a convergent sequence.

Let $(\Omega, \mathcal{F}, P)$ be a probability space. A G-valued *random variable* is a measurable function from $(\Omega, \mathcal{F})$ to $(G, \mathcal{B}(G))$. If X is such a random variable, its *law* or *distribution* is the measure $\mu_X \in \mathcal{P}(G)$ defined by $\mu_X(B) = P(X^{-1}(B))$ for all

$B \in \mathcal{B}(G)$. The product of two random variables X and Y is the random variable XY whose value at $\omega \in \Omega$ is $X(\omega)Y(\omega)$.[2] If X and Y are independent then the law of XY is the convolution $\mu_X * \mu_Y$.

4.2 The Fourier Transform of a Probability Measure

Let $\mathrm{Rep}(G)$ be the set of all unitary representations of G. So for each $\pi \in \mathrm{Rep}(G)$, $g \in G$, $\pi(g)$ acts as a unitary operator on the complex separable Hilbert space V_π. For each $\mu \in \mathcal{P}(G)$, we define its *Fourier transform* or *characteristic function* $\widehat{\mu}(\pi)$ at $\pi \in \mathrm{Rep}(G)$ to be the bounded linear operator on V_π defined as a Bochner integral (see e.g. Cohn [50] Appendix E, pp. 350–354) by:

$$\widehat{\mu}(\pi)\psi = \int_G \pi(g^{-1})\psi\mu(dg), \tag{4.2.3}$$

for each $\psi \in V_\pi$. Equivalently, it may be defined as a Pettis integral to be the unique bounded linear operator on V_π for which

$$\langle\widehat{\mu}(\pi)\phi, \psi\rangle = \int_G \langle\pi(g^{-1})\phi, \psi\rangle\mu(dg), \tag{4.2.4}$$

for all $\phi, \psi \in V_\pi$ (c.f. Heyer [93], Siebert [185] and Hewitt and Ross [92], pp. 77–87).

Note that if μ is absolutely continuous with respect to a given left Haar measure on μ and has density $f \in L^1(G)$, then our definition is such that $\widehat{\mu}(\pi) = \widehat{f}(\pi)$, where $\widehat{f}(\pi)$ is as defined in Chap. 2.[3]

From now until Sect. 4.7, *we will take G to be a compact Lie group and restrict π to be an irreducible representation.*[4]

So (observing our usual convention of identifying equivalence classes with representative elements) we will from now on always take $\pi \in \widehat{G}$. Then $\widehat{\mu}(\pi)$ is a $d_\pi \times d_\pi$ matrix and both (4.2.3) and (4.2.4) are equivalent to defining the matrix elements

$$\widehat{\mu}(\pi)_{ij} = \int_G \pi_{ij}(g^{-1})\mu(dg), \tag{4.2.5}$$

[2] If G is abelian, then the binary operation in the group is usually written additively.

[3] It is common in the literature to see the alternative definition "$\widehat{\mu}(\pi) = \int_G \pi(g)\mu(dg)$" which is natural for probabilists but which clashes with the analysts' convention that we introduced in Chap. 2.

[4] Many theorems that we state hold under more general conditions on G. The reader who wants minimal assumptions may consult the original sources, or check what is really needed from the proof.

for $1 \leq i, j \leq d_\pi$.

It is often convenient to write (4.2.5) using the simplified notation:

$$\widehat{\mu}(\pi) := \int_G \pi(g^{-1})\mu(dg).$$

Example 1 Dirac Mass. If $\mu = \delta_g$ for some $g \in G$, then it is easily verified that for all $\pi \in \widehat{G}$, $\widehat{\mu}(\pi) = \pi(g^{-1})$. In particular, $\widehat{\delta_e} = I_\pi$.

Example 2 Normalised Haar measure. We again denote this measure by m. It is easy to see that $m \in \mathcal{P}_{sc}(G)$. We have

$$\widehat{m}(\pi) = \begin{cases} 0 & \text{if } \pi \neq \pi_0 \\ 1 & \text{if } \pi = \pi_0 \end{cases}.$$

To see this, it is sufficient to observe that for all $\pi \in \widehat{G}$, $1 \leq i, j \leq d_\pi$,

$$\widehat{m}(\pi)_{ij} = \int_G \pi_{ij}(g^{-1})dg = \langle 1, \pi_{ij}\rangle_{L^2(G)},$$

and the result then follows by Peter-Weyl theory (Theorem 2.2.4).

Example 3 Standard Gaussian Measures. We recall the discussion of the heat kernel in Sect. 3.1.1. Now fix a parameter $\sigma > 0$ and consider the heat equation:

$$\frac{\partial u}{\partial t} = \sigma \Delta u. \tag{4.2.6}$$

We write the corresponding heat kernel as $k_\sigma \in C^\infty((0, \infty) \times G, \mathbb{R})$, and for fixed $t > 0$ we write $k_{t,\sigma}(\cdot) := k_\sigma(t, \cdot) \in C^\infty(G, \mathbb{R})$. Taking $f = 1$ in (3.1.8), we see immediately from (4.2.6) that $\int_G k_{t,\sigma}(g)dg = 1$, and so $k_{t,\sigma}$ is the density of a measure $\gamma_{t,\sigma} \in \mathcal{P}(G)$ which we call a *standard Gaussian measure* with parameter σ.[5] We now compute the Fourier transform. Using the smoothness of $t \to k_{t,\sigma}$ and dominated convergence, we deduce that for all $\pi \in \widehat{G}$,

$$\int_G \pi(g^{-1})\frac{\partial k_{t,\sigma}(g)}{\partial t}dg = \frac{\partial}{\partial t}\int_G \pi(g^{-1})k_{t,\sigma}(g)dg,$$

and so the mapping $t \to \widehat{k_{t,\sigma}}(\pi)$ is differentiable. Taking Fourier transforms of both sides of (4.2.6) then yields that

[5] If we were to take a strict analogy with the well-known theory in Euclidean space, we would only use the terminology "standard" Gaussian measure for the case where $\sigma t = \frac{1}{2}$.

$$\frac{\partial \widehat{k_{t,\sigma}}(\pi)}{\partial t} = \sigma \widehat{\Delta k_{t,\sigma}}(\pi)$$
$$= -\sigma \kappa_\pi \widehat{k_{t,\sigma}}(\pi).$$

Since $\widehat{k_{0,\sigma}}(\pi) = \widehat{\delta_e}(\pi) = I_\pi$, we deduce that

$$\widehat{k_{t,\sigma}}(\pi) = e^{-t\sigma\kappa_\pi} I_\pi. \tag{4.2.7}$$

The next theorem summarises some key properties of the Fourier transform (see also Heyer [93]):

Theorem 4.2.1 *For all $\mu, \mu_1, \mu_2 \in \mathcal{P}(G), \pi \in \widehat{G}$,*

1. $\widehat{\mu}(\pi_0) = 1$,
2. $\widehat{\mu_1 * \mu_2}(\pi) = \widehat{\mu_2}(\pi)\widehat{\mu_1}(\pi)$,
3. $||\widehat{\mu}(\pi)||_{op} \leq 1$,
4. $\widehat{\widetilde{\mu}}(\pi) = \widehat{\mu}(\pi)^*$.

Proof 1. is obvious.

2. For all $1 \leq i, j \leq d_\pi$

$$\begin{aligned}\widehat{\mu_1 * \mu_2}(\pi)_{ij} &= \int_G \pi_{ij}(h^{-1}g^{-1})\mu_1(dg)\mu_2(dh) \\ &= \sum_{k=1}^{d_\pi} \left(\int_G \pi_{ik}(h^{-1})\mu_2(dh)\right)\left(\int_G \pi_{kj}(g^{-1})\mu_1(dg)\right) \\ &= [\widehat{\mu_2}(\pi)\widehat{\mu_1}(\pi)]_{ij}.\end{aligned}$$

3. For all $\phi \in V_\pi$,

$$\begin{aligned}||\widehat{\mu}(\pi)\phi|| &= \left|\left|\int_G \pi(g^{-1})\phi\mu(dg)\right|\right| \\ &\leq \int_G ||\pi(g^{-1})\phi||\mu(dg) \\ &= ||\phi||.\end{aligned}$$

4.

$$\widehat{\widetilde{\mu}}(\pi) = \int_G \pi(g^{-1})\widetilde{\mu}(dg)$$
$$= \int_G \pi(g)\mu(dg)$$
$$= \left(\int_G \pi(g^{-1})\mu(dg)\right)^* = \widehat{\mu}(\pi)^*.$$ □

Corollary 4.2.1 *The measure $\mu \in \mathcal{P}_s(G)$ if and only if the matrix $\widehat{\mu}(\pi)$ is self-adjoint*[6] *for all $\pi \in \widehat{G}$.*

Proof Necessity is immediate from Theorem 4.2.1 (4). For sufficiency it is enough to observe that if the self-adjointness condition holds, then for all $\pi \in \widehat{G}$, $\phi, \psi \in V_\pi$,

$$\int_G \langle \pi(g)\phi, \psi\rangle \mu(dg) = \int_G \langle \pi(g)\phi, \psi\rangle \widetilde{\mu}(dg).$$

By linearity we find that $\int_G f(g)\mu(dg) = \int_G f(g)\widetilde{\mu}(dg)$ for all $f \in \mathcal{E}(G)$ which is norm dense in $C(G)$ by the Peter-Weyl theorem (Theorem 2.2.4). By extension of bounded linear functionals, we then see that $\int_G f(g)\mu(dg) = \int_G f(g)\widetilde{\mu}(dg)$ for all $f \in C(G)$, and the result follows from the Riesz representation theorem. □

The next theorem generalises Theorem 2.4.1 (see also Hewitt and Ross [92] Theorem 28.48, pp. 84–85).

Theorem 4.2.2 *The measure $\mu \in \mathcal{P}_c(G)$ if and only if $\widehat{\mu}(\pi) = c_\pi I_\pi$, where $c_\pi \in \mathbb{C}$, for all $\pi \in \widehat{G}$.*

Proof Necessity is established by Schur's lemma just as in Theorem 2.4.1. For sufficiency, for each $h \in G$ define $\mu^h \in \mathcal{P}(G)$ by $\mu^h(A) = \mu(hAh^{-1})$ for $A \in \mathcal{B}(G)$. Then arguing as in the proof of Theorem 2.4.1 we obtain for all $\pi \in \widehat{G}$, $\int_G \pi(g^{-1})\mu(dg) = \int_G \pi(g^{-1})\mu^h(dg)$, and so for all $\phi, \psi \in V_\pi$,

$$\int_G \langle \pi(g)\phi, \psi\rangle \mu(dg) = \int_G \langle \pi(g)\phi, \psi\rangle \mu^h(dg).$$

We can now reach our desired conclusion by proceeding as in the proof of Corollary 4.2.1 □

Corollary 4.2.2 *The measure $\mu \in \mathcal{P}_{sc}(G)$ if and only if $\widehat{\mu}(\pi) = c_\pi I_\pi$, where $c_\pi \in \mathbb{R}$, for all $\pi \in \widehat{G}$.*

---

[6] i.e. hermitian, if you prefer that terminology.

Proof This follows immediately from Corollary 4.2.1 and Theorem 4.2.2. □

For example we find by (4.2.7) that standard Gaussian measure is both central and symmetric.

The remaining results in this section were originally due to Kawada and Itô [114]. The first of these establishes the injectivity of the Fourier transform:

Theorem 4.2.3 *Let $\mu_1, \mu_2 \in \mathcal{P}(G)$. Then $\widehat{\mu_1}(\pi) = \widehat{\mu_2}(\pi)$ for all $\pi \in \widehat{G}$ if and only if $\mu_1 = \mu_2$.*

Proof Sufficiency is immediate. For necessity let $f \in C(G)$, and $\epsilon > 0$ be arbitrary. By the Peter-Weyl theorem (Theorem 2.2.4) there exists $\widehat{G}_0 \subset \widehat{G}$ with $\#\widehat{G}_0 \in \mathbb{N}$ such that

$$\sup_{g \in G}\left| f(g) - \sum_{\pi \in \widehat{G}_0} \sum_{i,j=1}^{d_\pi} \alpha_{ij}^{(\pi)} \pi_{ij}(g)\right| < \frac{\epsilon}{2},$$

where $\alpha_{ij}^{(\pi)} \in \mathbb{C}$ $(1 \leq i, j \leq d_\pi)$. Then for $k = 1, 2$ we find that

$$\left| \int_G f(g)\mu_k(dg) - \sum_{\pi \in \widehat{G}_0} \sum_{i,j=1}^{d_\pi} \alpha_{ij}^{(\pi)} \widehat{\mu}_k(\pi)_{ij}\right| < \frac{\epsilon}{2}.$$

But since $\widehat{\mu}_1(\pi)_{ij} = \widehat{\mu}_2(\pi)_{ij}$ for all $1 \leq i, j \leq d_\pi$, we deduce that

$$\left| \int_G f(g)\mu_1(dg) - \int_G f(g)\mu_2(dg)\right| < \epsilon,$$

and the result follows by the fact that ϵ is arbitrary and by use of the Riesz representation theorem. □

Theorem 4.2.4 *Let $\mu_1, \mu_2 \in \mathcal{P}(G)$. Then $\mu_1 * \mu_2 = \mu_2 * \mu_1$ if and only if $\widehat{\mu_1}(\pi)\widehat{\mu_2}(\pi) = \widehat{\mu_2}(\pi)\widehat{\mu_1}(\pi)$ for all $\pi \in \widehat{G}$.*

Proof Necessity follows immediately from Theorem 4.2.1(2). For sufficiency, observe that by Theorem 4.2.1(2) again

$$\widehat{\mu_1 * \mu_2}(\pi) = \widehat{\mu_2}(\pi)\widehat{\mu_1}(\pi) = \widehat{\mu_1}(\pi)\widehat{\mu_2}(\pi) = \widehat{\mu_2 * \mu_1}(\pi),$$

and then apply Theorem 4.2.3. □

Theorem 4.2.5 *Let $(\mu_n, n \in \mathbb{N})$ be a sequence of measures in $\mathcal{P}(G)$. Then $\mu_n \underset{w}{\to} \mu$ as $n \to \infty$ if and only if $\widehat{\mu_n}(\pi)_{ij} \to \widehat{\mu}(\pi)_{ij}$ as $n \to \infty$ for all $1 \leq i, j \leq d_\pi$, $\pi \in \widehat{G}$.*

Proof If $\mu_n \underset{w}{\to} \mu$ as $n \to \infty$, then

$$\lim_{n\to\infty} \widehat{\mu_n}(\pi)_{ij} = \lim_{n\to\infty} \int_G \pi_{ij}(g^{-1})\mu_n(dg) = \int_G \pi_{ij}(g^{-1})\mu(dg) = \widehat{\mu}(\pi)_{ij}.$$

Conversely, if $\widehat{\mu_n}(\pi)_{ij} \to \widehat{\mu}(\pi)_{ij}$ as $n \to \infty$ for all $1 \leq i, j \leq d_\pi, \pi \in \widehat{G}$, then using the same notation, and a similar argument to that given in the proof of Theorem 4.2.3, we first observe that for any $f \in C(G)$, $\epsilon > 0$ there exists $\widehat{G}_0 \subset \widehat{G}$ with $\#\widehat{G}_0 \in \mathbb{N}$ so that for all $n \in \mathbb{N}$

$$\left| \int_G f(g)\mu_n(dg) - \sum_{\pi\in\widehat{G}_0} \sum_{i,j=1}^{d_\pi} \alpha_{ij}^{(\pi)} \widehat{\mu_n}(\pi)_{ij} \right| < \frac{\epsilon}{3},$$

and also

$$\left| \int_G f(g)\mu(dg) - \sum_{\pi\in\widehat{G}_0} \sum_{i,j=1}^{d_\pi} \alpha_{ij}^{(\pi)} \widehat{\mu}(\pi)_{ij} \right| < \frac{\epsilon}{3}.$$

But we can also find $n_1 \in \mathbb{N}$ so that if $n > n_1$ we have

$$|\widehat{\mu_n}(\pi)_{ij} - \widehat{\mu}(\pi)_{ij}| < \frac{\epsilon}{3C}$$

for all $1 \leq i, j \leq d_\pi, \pi \in \widehat{G}_0$ where $C := \sum_{\pi\in\widehat{G}_0} \sum_{i,j=1}^{d_\pi} |\alpha_{ij}^{(\pi)}|$. From these estimates we deduce that for all $n > n_1$,

$$\left| \int_G f(g)\mu_n(dg) - \int_G f(g)\mu(dg) \right| < \epsilon,$$

and this gives the desired weak convergence. □

Then final result of this section gives a compact Lie group version of the celebrated *Lévy convergence theorem* for sequences of probability measures in Euclidean space.

Theorem 4.2.6 (Kawada,Itô,Lévy convergence theorem) *Suppose that* $(\mu_n, n \in \mathbb{N})$ *is a sequences of measures in* $\mathcal{P}(G)$ *and that there exists a family of compatible matrices* $(Y(\pi), \pi \in \widehat{G})$ *so that* $\widehat{\mu_n}(\pi)_{ij} \to Y(\pi)_{ij}$ *as* $n \to \infty$ *for all* $1 \leq i, j \leq d_\pi, \pi \in \widehat{G}$. *Then there exists* $\mu \in \mathcal{P}(G)$ *for which* $\mu_n \underset{w}{\to} \mu$ *as* $n \to \infty$ *and* $\widehat{\mu}(\pi) = Y(\pi)$ *for all* $\pi \in \widehat{G}$.

Proof Let $f \in C(G)$. Once again using (a straightforward variation of) the same notation to that used in the proof of Theorem 4.2.3, we can assert that given any $m \in \mathbb{N}$ there exists $\widehat{G}_0 \subset \widehat{G}$ with $\#\widehat{G}_0 \in \mathbb{N}$ so that

$$\sup_{g\in G}\left|f(g)-\sum_{\pi\in\widehat{G}_0}\sum_{i,j=1}^{d_\pi}\alpha_{ij}^{(\pi,m)}(f)\pi_{ij}(g)\right|<\frac{1}{2^m},$$

and so for all $n\in\mathbb{N}$

$$\left|\int_G f(g)\mu_n(dg)-\sum_{\pi\in\widehat{G}_0}\sum_{i,j=1}^{d_\pi}\alpha_{ij}^{(\pi,m)}(f)\widehat{\mu_n}(\pi)_{ij}\right|<\frac{1}{2^m}.$$

Now given any $\epsilon>0$ and choosing n sufficiently large, we obtain for such n and arbitrary m that:

$$\left|\int_G f(g)\mu_n(dg)-\sum_{\pi\in\widehat{G}_0}\sum_{i,j=1}^{d_\pi}\alpha_{ij}^{(\pi,m)}(f)Y(\pi)_{ij}\right|<\frac{1}{2^m}+\epsilon.$$

Define $\Gamma_m(f):=\sum_{\pi\in\widehat{G}_0}\sum_{i,j=1}^{d_\pi}\alpha_{ij}^{(\pi,m)}(f)Y(\pi)_{ij}$. Then from the last inequality we deduce that $(\Gamma_m(f),m\in\mathbb{N})$ is a Cauchy sequence, and hence convergent to $\Gamma(f)\in\mathbb{C}$. Again from the last inequality, we deduce that

$$\Gamma(f)=\lim_{n\to\infty}\int_G f(g)\mu_n(dg),$$

from which it follows that $f\to\Gamma(f)$ is a positive linear functional on $C(G)$ for which $\Gamma(1)=1$. Hence by the Riesz representation theorem, there exists a probability measure $\mu\in\mathcal{P}(G)$ for which

$$\Gamma(f)=\int_G f(g)\mu(dg),$$

for all $f\in C(G)$ and this gives the required weak convergence. The fact that $\widehat{\mu}(\pi)=Y(\pi)$ for all $\pi\in\widehat{G}$ then follows from Theorem 4.2.3. □

4.3 Lo-Ng Positivity

Let μ be a Borel probability measure defined on a locally compact abelian group G (with group composition written additively). Let $\widehat{G}$ be the (abelian) dual group of characters (see Sect. 2.2.2) and let the neutral element in $\widehat{G}$ be $\widehat{e}$. In this case we have $\widehat{\mu}(\chi)=\int_G\overline{\chi(g)}\mu(dg)$ for all $\chi\in\widehat{G}$. Let $F:\widehat{G}\to\mathbb{C}$. The celebrated *Bochner theorem* gives a necessary and sufficient condition for $F=\widehat{\mu}$, for some $\mu\in\mathcal{P}(G)$, and this is precisely that $F(\widehat{e})=1$, F is continuous at $\widehat{e}$ and F is positive definite, i.e.

$$\sum_{i,j=1}^{n} c_i \overline{c_j} F(x_i - x_j) \geq 0$$

for all $n \in \mathbb{N}, c_1, \ldots, c_n \in \mathbb{C}$ and all $x_1, \ldots, x_n \in \widehat{G}$ (see e.g. Heyer [96], pp. 162–184 or Rudin [171], pp. 19–21). There is an analogue of this result if G is a compact Lie group which we now describe. Further details and proofs are in Heyer [95], pp. 57–59.[7]

We recall the coefficient algebra $\mathcal{E}(G)$ of G from Chap. 2. We say that a linear functional $\phi : \mathcal{E}(G) \to \mathbb{C}$ is continuous if given any sequence $(f_n, n \in \mathbb{N})$ converging uniformly to $f \in \mathcal{E}(G)$, we have that $(\phi(f_n), n \in \mathbb{N})$ converges to $\phi(f)$. We say that ϕ is positive if $\phi(\overline{f} f) \geq 0$ for all $f \in \mathcal{E}(G)$.

Theorem 4.3.1 *If G is a compact Lie group, then for any positive continuous linear functional ϕ on $\mathcal{E}(G)$ for which $\phi(1) = 1$, there exists $\mu \in \mathcal{P}(G)$ so that*

$$\langle \widehat{\mu}(\pi)x, y \rangle = \phi(\langle \pi(\cdot)x, y \rangle),$$

for all $\pi \in \widehat{G}, x, y \in V_\pi$.

As an alternative to Bochner's theorem, we can find an interesting necessary and sufficient condition for a family of compatible matrices to be the Fourier transform of a finite measure if we introduce a new notion of positivity due to Lo and Ng [136], as we will now demonstrate. To this end let $C : \widehat{G} \to \mathcal{M}(\widehat{G})$ be a compatible mapping. We say that it is *Lo-Ng positive* if the following holds: Whenever $B : \widehat{G} \to \mathcal{M}(\widehat{G})$ is any other compatible mapping for which

$$\sum_{\pi \in S} d_\pi \mathrm{tr}(\pi(g) B(\pi)) \geq 0$$

for all $g \in G$ for some finite subset[8] S of $\widehat{G}$, then

$$\sum_{\pi \in S} d_\pi \mathrm{tr}(\pi(g) C(\pi) B(\pi)) \geq 0$$

for all $g \in G$. It is immediate that if C is Lo-Ng positive and $a \geq 0$, then aC is also Lo-Ng positive. The following gives a useful alternative criterion for Lo-Ng positivity:

Lemma 4.3.1 *The compatible mapping C is Lo-Ng positive if and only for all compatible mappings $B : \widehat{G} \to \mathcal{M}(\widehat{G})$, $\sum_{\pi \in S} d_\pi \mathrm{tr}(\pi(g) B(\pi)) \geq 0$ for all $g \in G$ for some finite subset S of $\widehat{G}$ implies that $\sum_{\pi \in S} d_\pi \mathrm{tr}(B(\pi) C(\pi)) \geq 0$.*

---

[7] As a result of reading an early version of this manuscript, Herbert Heyer [97] was inspired to prove a new Bochner-type theorem for *central* probability measures on compact groups.

[8] Our definition is slightly different from that of Lo and Ng, who introduce an ordering of the countable set $\widehat{G}$ and instead of taking arbitrary finite subsets of $\widehat{G}$ as we do, choose sets of the form $\{1, 2, \ldots, n\}$, with respect to their given ordering.

Proof First suppose that C is indeed Lo-Ng positive. Then the required result follows by taking $g = e$ in the definition. Conversely suppose the given condition holds on some finite subset S of $\widehat{G}$. By the assumption on B we have

$$\sum_{\pi \in S} d_\pi \mathrm{tr}(\pi(gh)B(\pi)) \geq 0$$

for all $g, h \in G$. It follows that

$$\sum_{\pi \in S} d_\pi \mathrm{tr}(\pi(g)(B(\pi)\pi(h))) \geq 0$$

for all $g \in G$. Then by the given condition, for all $h \in G$,

$$\sum_{\pi \in S} d_\pi \mathrm{tr}(\pi(h)C(\pi)B(\pi)) = \sum_{\pi \in S} d_\pi \mathrm{tr}(C(\pi)(B(\pi)\pi(h))) \geq 0,$$

and Lo-Ng positivity is established. □

Lemma 4.3.1 equips us with the tool to show that the set of all Lo-Ng positive compatible mappings is closed under taking adjoints. To be precise, let $C : \widehat{G} \to \mathcal{M}(\widehat{G})$ be a compatible mapping and define its adjoint $C^* : \widehat{G} \to \mathcal{M}(\widehat{G})$ by the prescription $C^*(\pi) := C(\pi)^*$ for all $\pi \in \widehat{G}$.

Lemma 4.3.2 *If C is a Lo-Ng positive compatible mapping, then so is C^*.*

Proof Let $B : \widehat{G} \to \mathcal{M}(\widehat{G})$ be a compatible mapping for which

$$\sum_{\pi \in S} d_\pi \mathrm{tr}(\pi(g)B(\pi)) \geq 0$$

for all $g \in G$ for some finite subset S of $\widehat{G}$. Then

$$\begin{aligned}\sum_{\pi \in S} d_\pi \mathrm{tr}(\pi(g)B(\pi)^*) &= \sum_{\pi \in S} d_\pi \mathrm{tr}(B(\pi)^*\pi(g)) \\ &= \sum_{\pi \in S} d_\pi \overline{\mathrm{tr}(\pi(g^{-1})B(\pi))} \\ &= \sum_{\pi \in S} d_\pi \mathrm{tr}(\pi(g^{-1})B(\pi)) \geq 0.\end{aligned}$$

So by Lemma 4.3.1,

$$\sum_{\pi \in S} d_\pi \mathrm{tr}(C(\pi)^*B(\pi)) = \sum_{\pi \in S} d_\pi \overline{\mathrm{tr}(B(\pi)^*C(\pi))} = \sum_{\pi \in S} d_\pi \mathrm{tr}(B(\pi)^*C(\pi)) \geq 0,$$

and the result follows. □

Before we proceed further we state a useful technical lemma

Lemma 4.3.3 *Let $B, C : \widehat{G} \to \mathcal{M}(\widehat{G})$ be compatible mappings and let S, S' be finite subsets of $\widehat{G}$ with $S' \subseteq S$. Then*

$$\int_G \left(\sum_{\pi' \in S'} d_{\pi'} \operatorname{tr}(\pi'(g^{-1}) B(\pi')) \right) \left(\sum_{\pi \in S} d_\pi \operatorname{tr}(\pi(g) C(\pi)) \right) dg \tag{4.3.8}$$
$$= \sum_{\pi \in S'} d_\pi \operatorname{tr}(B(\pi) C(\pi)).$$

Proof Write both traces on the left hand side of (4.3.8) as finite sums and then use the Schur orthogonality relations (Corollary 2.2.3). □

The next result begins to establish the link between Lo-Ng positivity and the Fourier transform. Let S be a finite subset of $\widehat{G}$ and $C : S \to \mathcal{M}(\widehat{G})$ be compatible (we may consider C as extended to the whole of $\widehat{G}$ by defining it to be the zero matrix on $\widehat{G} - S$). Note that $f_{S,C} \in C(G)$, where for each $g \in G$, $f_{S,C}(g) := \sum_{\pi \in S} d_\pi \operatorname{tr}(C(\pi)\pi(g))$.

Proposition 4.3.1 *Let S, C and $f_{S,C}$ be as above.*

1. *For all $\pi \in \widehat{G}$, $C(\pi) = \widehat{f_{S,C}}(\pi)$.*
2. *If $f_{S,C} \geq 0$, then C is Lo-Ng positive.*

Proof 1. This follows by uniqueness of Fourier coefficients in the Fourier expansion (2.3.7) of $f_{S,C}$.
2. Suppose that $B : \widehat{G} \to \mathcal{M}(\widehat{G})$ is a compatible mapping for which

$$\sum_{\pi \in S'} d_\pi \operatorname{tr}(\pi(g) B(\pi)) \geq 0$$

for all $g \in G$ and some finite subset S' of S. By the hypothesis on $f_{S,C}$ and (4.3.8), it follows that

$$\sum_{\pi \in S'} d_\pi \operatorname{tr}(B(\pi) C(\pi)) \geq 0$$

and so C is Lo-Ng positive by Lemma 4.3.1. □

Next we state another technical lemma:

Lemma 4.3.4 *There exists a sequence $(\psi_n, n \in \mathbb{N})$ of continuous non-negative functions on G, with each $\psi_n(g) = \sum_{\pi \in S_n} d_\pi z_\pi^{(n)} \chi_\pi(g)$ where S_n is a finite subset of $\widehat{G}$ and $z_\pi^{(n)} \in \mathbb{C}$ for all $\pi \in S_n$, $n \in \mathbb{N}$ which has the following properties:*

(i) $\int_G \psi_n(g) dg = 1$ *for all $n \in \mathbb{N}$;*
(ii) *Given any neighbourhood U of e and any $\epsilon > 0$ there exists $n_0 \in \mathbb{N}$ such that $\psi_n(g) < \epsilon$ for all $g \in U^c$ and all $n \geq n_0$,*

(iii) $\lim_{n\to\infty} z_\pi^{(n)} = 1$ *for all* $\pi \in \widehat{G}$.

Proof We follow Talman [202] Theorem A.7.1, pp. 96–98 for (i) and (ii) and Lo-Ng [136] for (iii).

(i) First note that if π is a finite-dimensional representation of G, then by (3.3.13) we easily deduce that

$$\sup_{g\in G} |\chi_\pi(g)| \leq d_\pi.$$

Next observe that since G is a compact Lie group, it has a faithful finite-dimensional representation π (see e.g. Theorem 4.1 in Bröcker and tom Dieck [36], pp. 136–137) and for all $g, h \in G$ with $g \neq h$,

$$\sum_{i,j=1}^{d_\pi} |\pi_{ij}(g) - \pi_{ij}(h)|^2 > 0,$$

indeed, if there were equality, π would not be injective. Now

$$\begin{aligned}
&\sum_{i,j=1}^{d_\pi} |\pi_{ij}(g) - \pi_{ij}(h)|^2 \\
&= \sum_{i,j=1}^{d_\pi} (\pi_{ij}(g) - \pi_{ij}(h))(\overline{\pi_{ij}(g) - \pi_{ij}(h)}) \\
&= \sum_{i,j=1}^{d_\pi} (\pi_{ij}(g) - \pi_{ij}(h))(\pi_{ji}(g^{-1}) - \pi_{ji}(h^{-1})) \\
&= 2\chi_\pi(e) - \chi_\pi(gh^{-1}) - \overline{\chi_\pi(gh^{-1})}.
\end{aligned}$$

Let $\pi' := \pi \oplus \overline{\pi}$. Then for all $g \in G$, $\chi_{\pi'}(g) = \chi_\pi(g) + \overline{\chi_\pi(g)}$, and we deduce from the last display that for all $g \in G \setminus \{e\}$

$$\chi_{\pi'}(g) < \chi_{\pi'}(e) = d_{\pi'}.$$

Incorporating this with our earlier estimate, we see that for all $g \in G \setminus \{e\}$

$$-d_{\pi'} \leq \chi_{\pi'}(g) < d_{\pi'}.$$

Now define a new representation π'' of G to be the direct sum of π' and $d_{\pi'}$ copies of the trivial representation. Then for all $g \in G$,

$$\chi_{\pi''}(g) = d_{\pi'} + \chi_{\pi'}(g),$$

and the estimate just established yields, for all $g \in G \setminus \{e\}$,

$$0 \leq \chi_{\pi''}(g) < 2d_{\pi'} = d_{\pi''}.$$

Now for each $n \in \mathbb{N}, g \in G$ define $\psi_n(g) := c_n \chi_{\pi''}(g)^n$, where $c_n := \left(\int_G \chi_{\pi''}(g)^n dg\right)^{-1}$. Then by construction ψ_n is continuous, non-negative and $\int_G \psi_n(g)dg = 1$. By Theorem 2.4.2 (ii), $\chi_{\pi''}(g)^n$ is the value at g of the character of the n-fold tensor product of π'', and so by Theorem 2.4.2 (iv), $\chi_{\pi''}(g)^n = \sum_{\pi \in S_n} m_\pi^{(n)} \chi_\pi$ (where $m_\pi^{(n)}$ is a non-negative integer). Hence the complex numbers $z_\pi^{(n)}$ appearing in the statement of the lemma are given by $z_\pi^{(n)} = \dfrac{c_n m_\pi^{(n)}}{d_\pi}$.

(ii) For simplicity we write $\chi := \chi_{\pi''}$ and $d := d_{\pi''}$ for the remainder of this proof. Let U be an open neighbourhood of e. Then $G \setminus U$ is compact, and so there exists $g_0 \in G \setminus U$ for which $\chi(g_0) = \sup_{g \in G \setminus U} \chi(g)$ and we have $\chi(g_0) < d$. By continuity of $g \to \chi(g)$ at $g = e$, given any $\varepsilon > 0$ there exists an open neighbourhood V of e so that if $g \in V$, then $d - \varepsilon < \chi(g) < d + \varepsilon$. Now choose $\varepsilon = \dfrac{d - \chi(g_0)}{2}$ and we see that for all $g \in G, \chi(g) > \dfrac{d + \chi(g_0)}{2}$. Consequently, for each $n \in \mathbb{N}$,

$$\int_V \chi(g)^n dg > m(V) \left(\frac{d + \chi(g_0)}{2}\right)^n.$$

Now

$$\begin{aligned} c_n &< \left(\int_V \chi(g)^n dg\right)^{-1} \\ &< \frac{1}{m(V)} \left(\frac{2}{d + \chi(g_0)}\right)^n. \end{aligned}$$

Then for all $g \in G \setminus U$ we have

$$\begin{aligned} \psi_n(g) &< \frac{1}{m(V)} \left(\frac{2\chi(g)}{d + \chi(g_0)}\right)^n \\ &\leq \frac{1}{m(V)} \left(\frac{2\chi(g_0)}{d + \chi(g_0)}\right)^n, \end{aligned}$$

and we can make the quantity on the right hand side arbitrarily small by taking n to be sufficiently large.

(iii) If we take the inner product in $L^2(G)$ of ψ_n with the character of an arbitrary representation in $\widehat{G}$ and use Theorem 2.4.3, we can easily deduce that for each $n \in \mathbb{N}, \pi \in S_n, z_\pi^{(n)} = \frac{1}{d_\pi} \int_G \psi_n(g) \chi_\pi(g^{-1}) dg$. Then we find that

$$
\begin{aligned}
|z_\pi^{(n)} - 1| &= \frac{1}{d_\pi}\left|\int_G \psi_n(g)\chi_\pi(g^{-1})dg - d_\pi \int_G \psi_n(g)dg\right| \\
&\leq \frac{1}{d_\pi}\int_{U^c} \psi_n(g)|\chi_\pi(g^{-1}) - d_\pi|dg + \frac{1}{d_\pi}\int_U \psi_n(g)|\chi_\pi(g^{-1}) - d_\pi|dg,
\end{aligned}
$$

and the required result follows by taking U sufficiently small, n sufficiently large, and using the result of (ii) and the fact that $g \to \chi_\pi(g^{-1})$ is continuous, and takes the value d_π at e. □

The next result is the main one of this section.

Theorem 4.3.2 (The Lo-Ng Criterion) *Let $C : \widehat{G} \to \mathcal{M}(\widehat{G})$ be compatible. Then $C(\pi) = \widehat{\mu}(\pi)$ for all $\pi \in \widehat{G}$ where $\mu \in \mathcal{P}(G)$ if and only if C is Lo-Ng positive with $C(\pi_0) = 1$. Furthermore, μ is the weak limit of a sequence $(\mu_n, n \in \mathbb{N})$, where for each $n \in \mathbb{N}$, $\mu_n \in \mathcal{P}(G)$ is absolutely continuous with respect to Haar measure and has Radon-Nikodym derivative*

$$
h_n(g) = \sum_{\pi \in S_n} z_\pi^{(n)} d_\pi \mathrm{tr}(\pi(g)C(\pi))
$$

for all $g \in G$, where $\#S_m < \#S_n < \infty$ if $m < n$ and $\pi_0 \in S_n$ for all $n \in \mathbb{N}$.

Proof Assume that $\mu \in \mathcal{P}(G)$ and $\sum_{\pi \in S} d_\pi \mathrm{tr}(\pi(g)B_\pi) \geq 0$ for all $g \in G$ and some finite subset S of $\widehat{G}$. Then

$$
\sum_{\pi \in S} d_\pi \mathrm{tr}(\widehat{\mu}(\pi)B_\pi) = \int_G \sum_{\pi \in S} d_\pi \mathrm{tr}(\pi(g^{-1})B_\pi)\mu(dg) \geq 0,
$$

and so $\widehat{\mu} : \widehat{G} \to \mathcal{M}(\widehat{G})$ is Lo-Ng positive.

Conversely (and using the notation of Lemma 4.3.4), we have that for all $n \in \mathbb{N}$, $\psi_n(g) = \sum_{\pi \in S_n} d_\pi \mathrm{tr}(\pi(g)[z_\pi^{(n)} I_{d_\pi}]) \geq 0$ by Lemma 4.3.4, and so since C is assumed to be Lo-Ng positive, it follows that

$$
\sum_{\pi \in S_n} d_\pi \mathrm{tr}(\pi(g) z_\pi^{(n)} C(\pi)) \geq 0.
$$

By Lemma 4.3.1 we deduce that the compatible mapping whose value at $\pi \in S_n$ is $z_\pi^{(n)} C(\pi)$ (and whose value at $\pi \notin S_n$ is the zero matrix) is also Lo-Ng positive. By Proposition 4.3.1 (1),

$$
z_\pi^{(n)} C(\pi) = \int_G \pi(g^{-1}) h_n(g) dg.
$$

Since h_n is continuous it is integrable, and as h_n is non-negative, we can define a Borel measure μ_n on G whose Radon-Nikodym derivative is h_n. Using Peter-Weyl theory (Corollary 2.2.4), we have

$$\begin{aligned}\mu_n(G) &= \int_G h_n(g)dg \\ &= \int_G h_n(g)\pi_0(g)dg \\ &= z_{\pi_0}^{(n)} C(\pi_0) = 1.\end{aligned}$$

The fact that $z_{\pi_0}^{(n)} = 1$ follows from Lemma 4.3.4 (i) and the formula $z_\pi^{(n)} = \frac{1}{d_\pi}\int_G \psi_n(g)\chi_\pi(g^{-1})dg$ that is established within the proof of that same lemma. By Prohorov's theorem, we can now assert that there is a subsequence $(\mu_{n_k}, k \in \mathbb{N})$ that converges weakly to a probability measure μ. By Theorem 4.2.5, we have $\lim_{k\to\infty} \widehat{\mu_{n_k}}(\pi) = \widehat{\mu}(\pi)$ for all $\pi \in \widehat{G}$. But by construction $\lim_{k\to\infty} \widehat{\mu_{n_k}}(\pi) = \lim_{k\to\infty} z_\pi^{(n_k)} C(\pi) = C(\pi)$ by Lemma 4.3.4 (iii). Hence the converse is established.

To prove the last part of the theorem let $h \in C(G)$. Then by the Peter-Weyl theorem (Theorem 2.2.4), there exists a sequence of matrices $(H_n, n \in \mathbb{N})$ where each H_n acts in a finite-dimensional complex Hilbert space of dimension d_n such that $h(g) = \lim_{n\to\infty} \sum_{i=1}^n d_i \mathrm{tr}(\pi_i(g)^* H_i)$, and the convergence is uniform in $g \in G$. Using Schur orthogonality and (4.3.8), we find that

$$\begin{aligned}&\int_G h(g)\left(\sum_{\pi\in S_n} d_\pi z_\pi^{(n)} \mathrm{tr}(\pi(g)C(\pi))\right) dg \\ &= \lim_{m\to\infty}\int_G \left(\sum_{i=1}^m d_i \mathrm{tr}(\pi_i(g)^* H_i)\right)\left(\sum_{\pi\in S_n} d_\pi z_\pi^{(n)} \mathrm{tr}(\pi(g)C(\pi))\right) dg \\ &= \sum_{\pi\in S_n} d_\pi \mathrm{tr}(H(\pi) z_\pi^{(n)} C(\pi)) \\ &= \int_G \left(\sum_{\pi\in S_n} d_\pi \mathrm{tr}(\pi(g)^* H(\pi))\right) h_n(g)dg \\ &\to \int_G h(g)\mu(dg),\end{aligned}$$

as $n \to \infty$, using the dominated convergence theorem.

Remark

1. Although Lo-Ng positivity is an interesting theoretical result, it seems very difficult to use in practice to determine whether a given family of compatible matrices really is the Fourier transform of a finite measure.
2. As positive-definiteness (in the usual sense) is a key component of Bochner's theorem on locally compact abelian groups, it is worth pointing out that there is a general notion of positive definiteness for functions on a more general locally compact group G. Indeed, a continuous function $f : G \to \mathbb{C}$ is positive definite if and only if $\sum_{i,j=1}^{n} c_i \overline{c_j} f(g_i g_j^{-1}) \geq 0$ for all $g_1, \ldots, g_n \in G, c_1, \ldots, c_n \in \mathbb{C}, n \in \mathbb{N}$. You can learn about these functions in e.g. Sect. 2.8 of Edwards [61] or section 32 of Hewitt and Ross [92]. Note that there is even a *Bochner theorem* which describes the structure of such functions as linear combinations of certain elementary ones, but readers should be warned that it is not related to the Bochner theorem that we discussed at the beginning of this section (i.e. it does not give information about Fourier transforms of finite measures).

4.4 Absolute Continuity

We investigate absolute continuity of probability measures on G with respect to normalised Haar measure m. We follow the account in Wehn [216].

Theorem 4.4.1 (Raikov-Williamson) *Let $\mu \in \mathcal{P}(G)$. Then $\mu \in \mathcal{P}_{ac}(G)$ if and only if either $\mu(Eg) \to \mu(E)$ or $\mu(gE) \to \mu(E)$ as $g \to e$ for all $E \in \mathcal{B}(G)$.*

Proof We only deal here with the case $\mu(Eg) \to \mu(E)$ as $g \to e$. The other limit is dealt with by a similar argument.

First assume that $\mu \ll m$ and let $f_\mu := \dfrac{d\mu}{dm}$. Then for all $E \in \mathcal{B}(G)$,

$$\begin{aligned} |\mu(Eg) - \mu(E)| &\leq \int_E |f_\mu(hg^{-1}) - f_\mu(h)| dh \\ &\leq ||R_{g^{-1}} f_\mu - f_\mu||_1 \to 0 \text{ as } g \to e, \end{aligned}$$

by Proposition 1.2.1. Conversely, suppose that $\mu(Eg^{-1}) \to \mu(E)$ as $g \to e$ and suppose that $E \in \mathcal{B}(G)$ exists with $m(E) = 0$ and $\mu(E) > 0$. We seek a contradiction. Let $\rho \in L^1(G)$ be such that $\rho \geq 0$ and $\int_G \rho(g)dg = 1$. Then we may define a measure $\nu_\rho \in \mathcal{P}_{ac}(G)$ by $\nu_\rho(A) = \int_A \rho(g)dg$ for all $A \in \mathcal{B}(G)$. For all $g \in G$,

$$\begin{aligned} \nu_\rho(g^{-1}E) &= \int_G \rho(h) 1_E(gh) dh \\ &= \int_G \rho(g^{-1}h) 1_E(h) dh = 0, \end{aligned}$$

since $m(E) = 0$. Hence by (4.1.1)

$$(\mu * \nu_\rho)(E) = \int_G \nu_\rho(g^{-1}E)\mu(dg) = 0.$$

But again by (4.1.1), we have

$$(\mu * \nu_\rho)(E) = \int_G \mu(Eg^{-1})\nu_\rho(dg) > 0,$$

and this yields the required contradiction. □

For each $\mu \in \mathcal{P}(G)$ we define the associated *convolution operator* $T_\mu : B_b(G) \to B_b(G)$ by

$$(T_\mu f)(\sigma) := (f * \mu)(\sigma) = \int_G f(\sigma\tau)\mu(d\tau),$$

for all $f \in B_b(G), \sigma \in G$. It is easy to see that T_μ is linear and a contraction. Furthermore, if $\mu, \nu \in \mathcal{P}(G)$ we have

$$T_{\mu*\nu} = T_\mu T_\nu. \tag{4.4.9}$$

It is an important fact that $T_\mu : C(G) \to C(G)$. To see this, let $\sigma_1, \sigma_2 \in G$ and observe that for all $f \in C(G)$,

$$|T_\mu f(\sigma_1) - T_\mu f(\sigma_2)| \leq \int_G |f(\sigma_1\tau) - f(\sigma_2\tau)|\mu(d\tau) \leq ||L_{\sigma_1^{-1}} f - L_{\sigma_2^{-1}} f||_\infty,$$

and the result follows by left uniform continuity of f (see Theorem A.2.1 in Appendix A.2).

Before we state and prove the next result we recall that a subset S of $C(G)$ is *equicontinuous* if given any $\epsilon > 0$, each $g \in G$ has an open neighbourhood U_g so that if $h \in U_g$, then $|f(g) - f(h)| < \epsilon$ for all $f \in S$. The next theorem was originally established by Raikov [164], and we follow the account of Wehn [216].

Theorem 4.4.2 (Raikov) *Let $\mu \in \mathcal{P}(G)$. Then $\mu \in \mathcal{P}_{ac}(G)$ if and only if $T_\mu : C(G) \to C(G)$ is compact.*

Proof First suppose that $\mu \ll m$ and write $\rho_\mu := \dfrac{d\mu}{dm}$. Let $(f_n, n \in \mathbb{N})$ be a bounded sequence in $C(G)$. Then for all $g, h \in G, n \in \mathbb{N}$,

$$T_\mu f_n(g) - T_\mu f_n(h) = \int_G f_n(g\tau)\rho_\mu(\tau)d\tau - \int_G f_n(h\tau)\rho_\mu(\tau)d\tau$$
$$= \int_G f_n(\tau)(\rho_\mu(g^{-1}\tau) - \rho_\mu(h^{-1}\tau))d\tau,$$

from which we easily deduce that

$$|T_\mu f_n(g) - T_\mu f_n(h)| \leq \sup_{n\in\mathbb{N}} ||f_n||_\infty ||L_g\rho_\mu - L_h\rho_\mu||_1.$$

Then equicontinuity of $\{T_\mu f_n, n \in \mathbb{N}\}$ follows from Proposition 1.2.1. Uniform boundedness of $\{T_\mu f_n, n \in \mathbb{N}\}$ is easily verified. We can now appeal to the Arzelà-Ascoli theorem to deduce that $\{T_\mu f_n, n \in \mathbb{N}\}$ is relatively compact, and so contains a convergent subsequence. It follows that T_μ is compact.

Conversely, suppose that T_μ is compact and let E be an open set in G. Then since 1_E is lower semi-continuous, we can find a sequence $(f_n, n \in \mathbb{N})$ in $C(G)$ which increases monotonically to 1_E (see e.g. Nagami [152]). So in particular this sequence is bounded. Hence by assumption, its image contains a convergent subsequence $(T_\mu f_{n_k}, k \in \mathbb{N})$, and (uniformly in) $g \in G$,

$$\lim_{k\to\infty} T_\mu f_{n_k}(g) = T_\mu 1_E(g) = \mu(g^{-1}E).$$

It follows that the mapping $g \to \mu(g^{-1}E)$ is continuous, and so $\mu(g^{-1}E) \to \mu(E)$ as $g \to e$. A similar argument holds for the case where E is compact. By regularity of μ, the same limiting behaviour holds for $E \in \mathcal{B}(G)$. So by Theorem 4.4.1 we deduce that $\mu \ll m$, as required. □

In the case $G = \Pi^1$, the celebrated theorem of F. and M. Riesz gives a sufficient condition for a probability measure μ to be absolutely continuous. In that case $\widehat{G} = \mathbb{Z}$ and

$$\widehat{\mu}(n) = \frac{1}{2\pi}\int_0^{2\pi} e^{-inx}\mu(dx)$$

for each $n \in \mathbb{Z}$. Their sufficient condition for absolute continuity is that $\widehat{\mu}(n) = 0$ for all $n < 0$ (see e.g. Katznelson [113] p. 113). This result has been extended to compact Lie groups by Brummelhuis [38] (see also [37]). For ease of exposition, we state it here in the case where G is also connected and semisimple. Let $\pi \in \widehat{G}$ and recall that $V_\pi = \bigoplus_{\mu\in\mathcal{W}(\pi)} V_\mu$ where $\mathcal{W}(\pi)$ is the set of weights of π. Let λ be the highest weight and define $V_\pi^0 := V_\pi \ominus V_\lambda$.

Theorem 4.4.3 (Brummelhuis) *Let G be a compact, connected, semisimple Lie group. If $\mu \in \mathcal{P}(G)$ is such that $\widehat{\mu}(\pi)v = 0$ for all $v \in V_\pi^0$ and for all $\pi \in \widehat{G}$, then $\mu \ll m$.*

4.5 Regularity of Densities

In this section, we will investigate conditions for a probability measure on a compact group to have a square-integrable, continuous and smooth density of various orders. Although this topic is closely related to that of the Sect. 4.4, we will make no use of the results that we obtained there.

In this section we first examine the case where $\mu \in \mathcal{P}(G)$ has a square-integrable density. The following result is established in Applebaum [8].

Theorem 4.5.1 *Let G be a compact Lie group. Then $\mu \in \mathcal{P}(G)$ has an L^2-density f_μ if and only if*

$$\sum_{\pi \in \widehat{G}} d_\pi ||\widehat{\mu}(\pi)||_{HS}^2 < \infty. \tag{4.5.10}$$

In this case

$$f_\mu = \sum_{\pi \in \widehat{G}} d_\pi \mathrm{tr}(\widehat{\mu}(\pi)\pi(\cdot)) \tag{4.5.11}$$

Proof For necessity, suppose that $f_\mu \in L^2(G)$ is the density of μ. Then $\widehat{f_\mu}(\pi) = \widehat{\mu}(\pi)$ for all $\pi \in \widehat{G}$ and (4.5.10) follows from the Parseval-Plancherel identity (2.3.8).

For sufficiency define

$$f_\mu := \sum_{\pi \in \widehat{G}} d_\pi \mathrm{tr}(\widehat{\mu}(\pi)\pi).$$

Then $f_\mu \in L^2(G)$ since by (2.3.8),

$$||f_\mu||_2^2 = \sum_{\pi \in \widehat{G}} d_\pi ||\widehat{\mu}(\pi)||_{HS}^2 < \infty,$$

and by uniqueness of Fourier coefficients (in the Hilbert space sense) $\widehat{f_\mu}(\pi) = \widehat{\mu}(\pi)$ for all $\pi \in \widehat{G}$.[9]

Since Haar measure is finite, $L^2(G) \subseteq L^1(G)$, and so $f_\mu \in L^1(G)$. Recall that by Theorem 2.2.4 $\mathcal{E}(G)$, which is the algebra of all continuous functions on G that have only finitely many non-zero Fourier coefficients, is norm dense in $C(G)$. Let $h \in \mathcal{E}(G)$. Then there exists a finite subset S of $\widehat{G}$ so that

$$h(\sigma) = \sum_{\pi \in S} d_\pi \mathrm{tr}(\widehat{h}(\pi)\pi(\sigma))$$

for all $\sigma \in G$. Furthermore, by the Schur orthogonality relations, $\widehat{h}(\pi) = 0$ if $\pi \in S^c$. Using the Parseval-Plancherel identity (2.3.9), for each $h \in \mathcal{E}(G)$:

[9] To verify this directly, compute $\langle f_\mu, \pi'_{ij}\rangle$ for each $\pi' \in \widehat{G}$, $1 \leq i, j \leq d'$; it is then a straightforward application of the Peter-Weyl theorem (Corollary 2.2.4) to deduce that $\widehat{f_\mu}(\pi')_{ij} = \widehat{\mu}(\pi')_{ij}$.

$$\begin{aligned}\int_G h(\sigma)\overline{f_\mu(\sigma)}d\sigma &= \sum_{\pi\in\widehat{G}} d_\pi \mathrm{tr}(\widehat{h}(\pi)\widehat{\mu}(\pi)^*)\\ &= \sum_{\pi\in S} d_\pi \mathrm{tr}(\widehat{h}(\pi)\widehat{\mu}(\pi)^*)\\ &= \int_G \sum_{\pi\in S} d_\pi \mathrm{tr}(\widehat{h}(\pi)\pi(\sigma))\mu(d\sigma)\\ &= \int_G h(\sigma)\mu(d\sigma).\end{aligned}$$

By a standard density argument, it then follows that

$$\int_G h(\sigma)\overline{f_\mu(\sigma)}d\sigma = \int_G h(\sigma)\mu(d\sigma),$$

for all $h \in C(G)$. The Riesz representation theorem implies that f_μ is real valued and $f_\mu(\sigma)d\sigma = \mu(d\sigma)$. The fact that f_μ is non-negative a.e. then follows from the Jordan decomposition for signed measures (see Appendix A.5). □

Note that we can also write (4.5.11) as

$$f_\mu = 1 + \sum_{\pi\in\widehat{G}\setminus\{\pi_0\}} d_\pi \mathrm{tr}(\widehat{\mu}(\pi)\pi(\cdot)),$$

and that such a representation is often found in the literature.

It is easily seen that if μ is central, so that $\widehat{\mu}(\pi) = c_\pi I_\pi$ for all $\pi \in \widehat{G}$ (where $c_\pi \in \mathbb{C}$), and has L^2-density f_μ, then f_μ is central (a.e.). Then from (4.5.11) we have

$$f_\mu = \sum_{\pi\in\widehat{G}} d_\pi c_\pi \chi_\pi, \tag{4.5.12}$$

in the L^2 sense.

Next we examine continuity of densities:

Proposition 4.5.1 *Let $\mu \in \mathcal{P}(G)$. A sufficient condition for μ to have a continuous density f_μ is that the infinite series $\sum_{\pi\in\widehat{G}} d_\pi \mathrm{tr}(\widehat{\mu}(\pi)\pi(\sigma))$ converges uniformly in $\sigma \in G$.*

Proof Define $f_\mu(\sigma) = \sum_{\pi\in\widehat{G}} d_\pi \mathrm{tr}(\widehat{\mu}(\pi)\pi(\sigma))$ for all $\sigma \in G$. Then $f_\mu \in C(G)$, and by uniqueness of Fourier coefficients, $\widehat{f_\mu}(\pi) = \widehat{\mu}(\pi)$ for all $\pi \in \widehat{G}$. Now argue as in the proof of Theorem 4.5.1. □

More concrete sufficient conditions for μ to have a continuous density are as follows. In the second of these, for each $\mu \in \mathcal{P}(G)$ we employ the notation $\widehat{\mu}(\lambda) := \widehat{\mu}(\pi_\lambda)$, where $\lambda \in D$ is the highest weight corresponding to $\pi_\lambda \in \widehat{G}$:

- $\sum_{\pi\in\widehat{G}} d_\pi^{\frac{3}{2}} ||\widehat{\mu}(\pi)||_{HS} < \infty$,
- $||\widehat{\mu}(\pi_\lambda)||_{HS} = O(|\lambda|^{-s})$ as $|\lambda| \to \infty$, where $s > r + \dfrac{m}{2}$.

The first of these is implicit in the first part of the proof of Proposition 3.3.2 (see also Faraut [63], pp. 117–119). and the second is a direct consequence of the statement of that same proposition.

Next we investigate differentiability of densities. Recall that $\{\kappa_\pi, \pi \in \widehat{G}\}$ is the Casimir spectrum of G.

Theorem 4.5.2 *If $\mu \in \mathcal{P}(G)$ and there exists $p \in \mathbb{N}$ so that*

$$\sum_{\pi\in\widehat{G}} d_\pi (1+\kappa_\pi)^p ||\widehat{\mu}(\pi)||_{HS}^2 < \infty,$$

then μ has a C^k density for all $k < p - \dfrac{d}{2}$.

Proof Since $\kappa_\pi \geq 0$ for all $\pi \in \widehat{G}$, we have $\sum_{\pi\in\widehat{G}} d_\pi ||\widehat{\mu}(\pi)||_{HS}^2 < \infty$, and so by Theorem 4.5.1, μ has a L^2-density f_μ and $\widehat{f_\mu}(\pi) = \widehat{\mu}(\pi)$ for all $\pi \in \widehat{G}$. The result then follows by Proposition 3.1.4, and the Sobolev embedding theorem (Theorem 3.1.3). □

The next result establishes necessary and sufficient conditions for densities to exist and be C^∞. It was first established in Applebaum [12]. Recall that $\mathcal{S}(D)$ is the Sugiura space that was introduced in Sect. 3.4.

Theorem 4.5.3 *For G a compact connected Lie group, $\mu \in \mathcal{P}(G)$ has a C^∞ density if and only if $\widehat{\mu} \in \mathcal{S}(D)$.*

Proof Necessity is obvious. For sufficiency it is enough by Theorem 3.4.3 to show that μ has an L^2-density. Choose $s > r$, so that Suguira's zeta function (see below) converges (c.f. Theorem 3.2.1). Then using Theorem 4.5.1 we have

$$\begin{aligned}\sum_{\lambda\in D-\{0\}} d_\lambda ||\widehat{\mu}_\lambda||_{HS}^2 &\leq N \sum_{\lambda\in D-\{0\}} |\lambda|^m ||\widehat{\mu}_\lambda||_{HS}^2\\ &\leq N \sup_{\lambda\in D-\{0\}} |\lambda|^{m+s} ||\widehat{\mu}_\lambda||_{HS}^2 \sum_{\lambda\in D-\{0\}} \frac{1}{|\lambda|^s}\\ &< \infty.\end{aligned}$$

□

The following result gives an application of Theorem 4.5.3. First we note a useful and easily verified inequality for matrices. If $A, B \in M_n(\mathbb{C})$, then

$$||AB||_{HS} \leq ||A||_{\text{op}} ||B||_{HS} \text{ and } ||AB||_{HS} \leq ||B||_{\text{op}} ||A||_{HS} \tag{4.5.13}$$

Corollary 4.5.1 *Let G be a compact connected Lie group. Let $\mu \in \mathcal{P}(G)$ be arbitrary and $\gamma_{t,\sigma}$ be a standard Gaussian measure with parameters $t, \sigma > 0$. Then the measures $\mu * \gamma_{t,\sigma}$ and $\gamma_{t,\sigma} * \mu$ have smooth densities.*

Proof It suffices to establish the result for $\mu * \gamma_{t,\sigma}$. First note that by Theorem 4.2.1, (4.5.13) and (4.2.7) for all $\lambda \in D$,

$$||\widehat{\mu * \gamma_{t,\sigma}}(\lambda)||_{HS} = ||\widehat{\gamma_{t,\sigma}}(\lambda)\widehat{\mu}(\lambda)||_{HS} \leq ||\widehat{\mu}(\lambda)||_{\text{op}}||\widehat{\gamma_{t,\sigma}}(\lambda)||_{HS} \leq d_\lambda e^{-t\sigma\kappa_\lambda}.$$

But then using the dimension estimate of Corollary 2.5.2 and (2.5.23) we obtain for all $p \in \mathbb{N}$,

$$\limsup_{|\lambda|\to\infty} |\lambda|^p ||\widehat{\mu * \gamma_{t,\sigma}}(\lambda)||_{HS} \leq \lim_{|\lambda|\to\infty} |\lambda|^p d_\lambda e^{-t\sigma\kappa_\lambda} \leq C \lim_{|\lambda|\to\infty} |\lambda|^{p+m} e^{-t\sigma|\lambda|^2} = 0,$$

and the result follows from Theorem 4.5.3. □

4.6 Idempotents and Convolution Powers

We say that $\mu \in \mathcal{P}(G)$ is *idempotent* if $\mu * \mu = \mu$. Equivalently by Theorems 4.2.1 and 4.2.3, μ is idempotent if and only if $\widehat{\mu}(\pi)^2 = \widehat{\mu}(\pi)$ for all $\pi \in \widehat{G}$. It is easy to see that normalised Haar measure m on G is idempotent. More generally, let H be a closed subgroup of G and let $m_H^{(0)}$ denote its normalised Haar measure. We extend $m_H^{(0)}$ to a measure $m_H \in \mathcal{P}(G)$ that has support H by the prescription

$$m_H(B) = m_H^{(0)}(B \cap H)$$

for all $B \in \mathcal{B}(G)$. For example if $H = \{e\}$, then $m_H = \delta_e$. It is again easy to see that m_H is always idempotent. The following result is due to Wendel [217] for compact groups, but note that it also holds in general locally compact groups (see Heyer [95] Theorem 1.2.10, p. 34).[10]

Theorem 4.6.1 *If $\mu \in \mathcal{P}(G)$ is idempotent, then $\mu = m_H$ for some closed subgroup H of G. Moreover, $H = supp(\mu)$.*

Proof For $H := \text{supp}(\mu)$, by (4.1.2) we have $H = H^2$, and so H is a semigroup under the group law. It is also closed, and hence compact. It is known that any subset of G that has these properties is a subgroup (see e.g. Lemma 2 in Gelbaum et. al. [71] and also Corollary 1.2.9. on p. 34 of Heyer [95]). Now let $f \in C(H, \mathbb{R})$, and define for each $h \in H$:

$$A_f(h) = \int_G f(gh)\mu(dg).$$

Using Proposition 1.2.1 it is easily verified that $A_f \in C(H, \mathbb{R})$. Now let h_0 be the point in G where A_f attains its maximum value. From now on we denote $f_1 := R_{h_o} f$.

[10] Our standing hypothesis remains that G is compact Lie, but observe that the proof of Theorem 4.6.1 requires no use of Lie structure.

Then A_{f_1} attains its maximum at e. Then since μ is idempotent:

$$\begin{aligned}A_{f_1}(e) &= \int_H f_1(g)\mu(dg)\\ &= \int_H f_1(g)(\mu * \mu)(dg)\\ &= \int_H\int_H f_1(g_1g_2)\mu(dg_1)\mu(dg_2)\\ &= \int_H A_{f_1}(g_2)\mu(dg_2).\end{aligned}$$

Hence we see that $\int_H (A_{f_1}(e) - A_{f_1}(g_2))\mu(dg_2) = 0$, and so $A_{f_1}(e) - A_{f_1}(g)$ for all $g \in H$. It follows by uniqueness of Haar measure, and the fact that $\mu(H) = 1$, that $\mu = m_H$, as required. □

Let $\mu \in \mathcal{P}(G)$ and $n \in \mathbb{N}$. We define the nth convolution power of μ to be $\mu^{*(n)} = \mu * \cdots * \mu$ (n times). Note that we then have for all $\pi \in \widehat{G}$, $\widehat{\mu^{*(n)}}(\pi) = \widehat{\mu}(\pi)^n$. Let $(\Omega, \mathcal{F}, P)$ be a probability space and $(X_n, n \in \mathbb{N})$ be a sequence of independent, identically distributed (or i.i.d.) G–valued random variables. Let $(S_n, n \in \mathbb{N})$ be the associated G-valued *random walk*, so that for each $n \in \mathbb{N}$, $S_n = X_1 X_2 \ldots X_n$. Then the law of S_n is precisely $\mu^{*(n)}$. It is of interest to study the asymptotic behaviour of the random walk as for large n. In particular we might consider the weak limit of $\mu^{*(n)}$ as $n \to \infty$. It is clear that if the limit exists, it is an idempotent, and so by Theorem 4.6.1

$$\lim_{n\to\infty} \mu^{*(n)} = m_H,$$

for some closed subgroup H of G.

Necessary and sufficient conditions for the limit to exist were found by Stromberg [198]. We quote his result but omit the proof (see also Heyer [95] Theorem 2.1.4, pp. 91–92).

Theorem 4.6.2 *Let $\mu \in \mathcal{P}(G)$ and let K be the smallest closed subgroup of G containing supp(μ). Then* $\lim_{n\to\infty} \mu^{*(n)}$ *exists if and only if supp(μ) is not contained in any coset of a proper closed normal subgroup of K.*

Kawada and Itô [114] established an *equidistribution theorem* which gives conditions for the limit to exist and be normalised Haar measure m on the whole group. First we need a definition. We say that $\mu \in \mathcal{P}(G)$ is *aperiodic* if supp(μ) is not contained in a left or right coset of a proper closed normal subgroup of G. We then have the following:

Theorem 4.6.3 (Kawada-Itô equidistribution theorem) *If $\mu \in \mathcal{P}(G)$ is aperiodic, then $(\mu^{*(n)}, n \in \mathbb{N})$ converges weakly to normalised Haar measure.*

Proof This is based on the proof of Theorem 8 in [114]. By Theorem 4.2.6, it is sufficient to show that $\lim_{n\to\infty}\widehat{\mu}(\pi)^n = 0$ for all non-trivial $\pi \in \widehat{G}$. This is clearly equivalent to the requirement that all the eigenvalues of $\widehat{\mu}$ have modulus strictly less than 1. Note that since $\widehat{\mu}(\pi)$ is a contraction, its eigenvalues cannot have moduli that exceed 1. Now let λ be an eigenvalue of $\widehat{\mu}(\pi)$. Then we can find a unitary matrix U_π acting in V_π so that

$$U_\pi\widehat{\mu}(\pi)U_\pi^{-1} = \begin{pmatrix} \lambda & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & & & & & \\ \cdot & & & & & \\ \cdot & & & D_\pi & & \\ \cdot & & & & & \\ 0 & & & & & \end{pmatrix},$$

where D_π is some $(d_\pi - 1) \times (d_\pi - 1)$ matrix. In particular, we have

$$\lambda = (U_\pi\widehat{\mu}(\pi)U_\pi^{-1})_{11} = \int_G (U_\pi\pi(g^{-1})U_\pi^{-1})_{11}\mu(dg),$$

and if $|\lambda| = 1$, $(U_\pi\pi(g^{-1})U_\pi^{-1})_{11} = \lambda$ for all $g \in G$ for which $g^{-1} \in \operatorname{supp}(\mu)$.

Now suppose that $\lambda = 1$. Then we have

$$\operatorname{supp}(\mu) \subseteq H := \left\{ g \in G, U_\pi\pi(g^{-1})U_\pi^{-1} = \begin{pmatrix} 1 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & & & & & \\ \cdot & & & & & \\ \cdot & & & E_\pi(g) & & \\ \cdot & & & & & \\ 0 & & & & & \end{pmatrix} \right\},$$

where $E_\pi(g)$ is some $(d_\pi - 1) \times (d_\pi - 1)$ matrix. But H is a proper closed subgroup of G and this contradicts aperiodicity of μ.

Now suppose that $\lambda = e^{i\theta}$ for some $\theta \in \mathbb{R}\backslash 2\pi\mathbb{Z}$. Then arguing as above we can find a unitary matrix U_π so that $e^{i\theta} = (U_\pi\widehat{\mu}(\pi)U_\pi^{-1})_{11} = \int_G (U_\pi\pi(g^{-1})U_\pi^{-1})_{11}\mu(dg)$ and

$$\operatorname{supp}(\mu) \subseteq \Gamma := \left\{ g \in G, U_\pi\pi(g^{-1})U_\pi^{-1} = \begin{pmatrix} e^{i\theta} & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & & & & & \\ \cdot & & & & & \\ \cdot & & & F_\pi(g) & & \\ \cdot & & & & & \\ 0 & & & & & \end{pmatrix} \right\},$$

where $F_\pi(g)$ is some $(d_\pi - 1) \times (d_\pi - 1)$ matrix. Now let $g_0 \in G$ be such that $(V_\pi \pi(g_0^{-1}) V_\pi^{-1})_{11} = e^{i\theta}$. Then it is easily verified that $\Gamma = g_0 H$, and this again contradicts aperiodicity. The required result follows. □

We briefly draw the reader's attention to more recent work in this area. Shlosman and Major [142] and Shlosman [179, 180] were able to extend the Kawada-Itô theorem to the case where μ has a density and there is uniform convergence of its convolution powers to the uniform density. Johnson and Suhov [109] used the Kullback-Liebler distance to obtain exponential rates of convergence and Harremoës [81] examined this from the perspective of uniform convergence of the rate distortion function.

4.7 Convolution Operators

In this and the Sect. 4.8 we will drop the condition that G be a compact Lie group. We work more generally and assume (unless otherwise stated) that G is a locally compact, Hausdorff and second countable topological group. Convolution operators were already introduced in Sect. 4.4 for compact Lie groups. Now we study them more systematically. Let $\mu \in \mathcal{P}(G)$. The associated *right convolution operator* $P_\mu^{(R)}$ is defined on $B_b(G)$ by the prescription $P_\mu^{(R)} f = f * \mu$ for $f \in B_b(G)$, so that

$$(P_\mu^{(R)} f)(g) = \int_G f(gh)\mu(dh)$$

for all $g \in G$. Similarly the *left convolution operator* $P_\mu^{(L)}$ is defined by

$$(P_\mu^{(L)} f)(g) = \int_G f(hg)\mu(dh).$$

The reader may check that these are related by the identity

$$P_\mu^{(L)} f = \widetilde{P_{\widetilde{\mu}}^{(R)} \widetilde{f}}.$$

Furthermore, if $\mu_1, \mu_2 \in \mathcal{P}(G)$, then $P_{\mu_1}^{(R)} P_{\mu_2}^{(L)} = P_{\mu_2}^{(L)} P_{\mu_1}^{(R)}$. We also have that for all $\mu \in \mathcal{P}(G)$, $g \in G$, $L_g P_\mu^{(R)} = P_\mu^{(R)} L_g$ and $R_g P_\mu^{(L)} = P_\mu^{(L)} R_g$.

From now on we will almost always work with $P_\mu^{(R)}$, which we will denote simply as P_μ.

Proposition 4.7.1 *For each $\mu \in \mathcal{P}(G)$,*

1. *P_μ is linear,*

2. *$P_\mu 1 = 1$,*
3. *If $f \geq 0$, then $P_\mu f \geq 0$,*
4. *P_μ is a contraction,*
5. *$P_\mu : C_0(G) \to C_0(G)$,*
6. *If $\mu_1, \mu_2 \in \mathcal{P}(G)$, then*

$$P_{\mu_1 * \mu_2} = P_{\mu_1} P_{\mu_2}.$$

Proof (1)–(4) are all easy exercises. For (5) first note that continuity is straightforward to establish by using a dominated convergence argument (or see the discussion before Theorem 4.4.2). To see that if f vanishes at infinity then so does $P_\mu f$, let $G_\infty := G \cup \{\infty\}$ be the one-point compactification of G (see Appendix A.1). We extend μ to a probability measure on $(G_\infty, \mathcal{B}(G_\infty))$ by the requirement that $\mu(\{\infty\}) = 0$. Then P_μ extends to a linear operator from $C(G_\infty)$ to $C(G_\infty)$. If $F \in C(G_\infty)$ then $F \in C_0(G)$ if and only if $\lim_{g \to \infty} F(g) = 0$. Since $R_h : C_0(G) \to C_0(G)$ for all $h \in G$, the required result follows by using the dominated convergence theorem. To show (6), let $f \in B_b(G)$ and $g \in G$. By Fubini's theorem,

$$\begin{aligned}
(P_{\mu_1 * \mu_2} f)(g) &= \int_G f(gh)(\mu_1 * \mu_2)(dh) \\
&= \int_G \int_G f(gh_1h_2)\mu_1(dh_1)\mu_2(dh_2) \\
&= \int_G \left(\int_G f(gh_1h_2)\mu_2(dh_2) \right) \mu_1(dh_1) \\
&= \int_G (P_{\mu_2} f)(gh_1)\mu_1(dh_1) \\
&= (P_{\mu_1} P_{\mu_2} f)(g). \qquad \square
\end{aligned}$$

Note that if G is compact, then we use $C(G)$ in place of $C_0(G)$ in Proposition 4.7.1(5).

There is a useful link between the convolution operator and Fourier transform of the measure on compact groups.

Proposition 4.7.2 *If μ is a Borel probability measure on a compact Lie group G, then for all $\pi \in \widehat{G}$, $1 \leq i, j \leq d_\pi$*

$$\widehat{\widehat{\mu}}(\pi)_{ij} = P_\mu \pi_{ij}(e).$$

Proof

$$\begin{aligned} P_\mu \pi_{ij}(e) &= \int_G \pi_{ij}(\tau)\mu(d\tau) \\ &= \int_G \pi_{ij}(\tau^{-1})\widetilde{\mu}(d\tau) \\ &= \widehat{\widetilde{\mu}}(\pi)_{ij} \end{aligned}$$ □

Returning to the general case, we easily deduce from Proposition 4.7.1 (6) that for all $\mu \in \mathcal{P}(G), n \in \mathbb{N}$,

$$P_{\mu^{*(n)}} = P_\mu^n. \tag{4.7.14}$$

Note also that P_μ^n is the n-step transition operator for the random walk $(S(n), n \in \mathbb{N})$. Indeed, for all $f \in B_b(G), g \in G$,

$$(P_\mu^n f)(g) = \mathbb{E}(f(gS(n))).$$

In the following we fix a right Haar measure m_R on G and consider the p-norms $||\cdot||_p$ in the spaces $L^p(G, m_R)$ for $1 \le p < \infty$.

Proposition 4.7.3 *For all $\mu \in \mathcal{P}(G)$, $f \in C_c(G)$, $1 \le p < \infty$,*

$$||P_\mu f||_p \le ||f||_p.$$

Proof By Jensen's inequality and Fubini's theorem,

$$\begin{aligned} ||P_\mu f||_p^p &= \int_G |P_\mu f(g)|^p m_R(dg) \\ &= \int_G \left| \int_G f(gh)\mu(dh) \right|^p m_R(dg) \\ &\le \int_G \int_G |f(gh)|^p \mu(dh) m_R(dg) \\ &= \int_G \int_G |f(gh)|^p m_R(dg)\mu(dh) \\ &= \int_G |f(g)|^p m_R(dg) = ||f||_p^p. \end{aligned}$$ □

We have just shown that P_μ is a contraction from $C_c(G)$ to $L^p(G, m_R)$ and so it extends to a contraction on the whole of $L^p(G, m_R)$. We will continue to denote the operator by the same symbol P_μ whenever we consider it as acting in the L^p-spaces. For the case $p = 2$ the reader may verify the useful result

$$P_\mu^* = P_{\widetilde{\mu}}. \tag{4.7.15}$$

Let f be a non-negative measurable function defined on G. It is said to be μ-*harmonic* if $P_\mu f = f$, μ-*superharmonic* if $P_\mu f \leq f$ and μ-*subharmonic* if $P_\mu f \geq f$. Clearly non-negative constant functions are μ-harmonic. In the Sect. 4.8 we will investigate a condition under which they are the only ones. Note that if $(S_n, n \in \mathbb{N})$ is the random walk associated to μ and f is bounded and μ-superharmonic, then $(f(S_n), n \in \mathbb{N})$ is a supermartingale with respect to the filtration $(\mathcal{F}_n, n \in \mathbb{N})$ of $\mathcal{F}$ where for each $n \in \mathbb{N}, \mathcal{F}_n := \sigma\{X_1, \ldots, X_n\}$. To see this, observe that by the Markov property

$$\mathbb{E}(f(S_{n+1})|\mathcal{F}_n) = (P_\mu f)(S_n) \leq f(S_n).$$

Similarly we obtain a martingale (respectively, submartingale) if f is μ-harmonic (respectively, μ-subharmonic).

Now let $\mathcal{M}(G)$ be the space of all regular Borel measures on G. For each $\mu \in \mathcal{M}(G)$, we may consider the dual action P_μ^* of P_μ on $\mathcal{M}(G)$ defined by the prescription:

$$(P_\mu^* \rho)(f) = \rho(P_\mu f),$$

for all $\rho \in \mathcal{M}(G)$, $f \in C_c(G)$. We generalise the definitions we gave above for functions and say that $\rho \in \mathcal{M}(G)$ is μ-harmonic if $P_\mu^* \rho = \rho$, μ-superharmonic if $P_\mu^* \rho \leq \rho$ and μ-subharmonic if $P_\mu^* \rho \geq \rho$.[11] Now let $\rho \in \mathcal{M}(G)$ be absolutely continuous with respect to m_R with $f_\rho := \dfrac{d\rho}{dm_R}$. It is left as an exercise for the reader to check that f_ρ is μ-superharmonic if and only if ρ is $\widetilde{\mu}$-superharmonic.

We give some useful properties of μ-superharmonic functions.

Proposition 4.7.4 *Let $\mu \in \mathcal{P}(G)$*

1. *Suppose that $\lambda, \rho \in \mathcal{M}(G)$. If ρ is μ-superharmonic, then so is $\lambda * \rho$.*
2. *If f is μ-superharmonic, then so is $f \wedge c$ for any $c > 0$.*

Proof

1. For all $f \in C_c(G)_+$ we have

[11] If $\rho_1, \rho_2 \in \mathcal{M}(G)$ we write $\rho_1 \geq \rho_2$ if $\rho_1(f) \geq \rho_2(f)$ for all $f \in C_c(G)_+$.

$$
\begin{aligned}
P_\mu^*(\lambda * \rho)(f) &= \int_G \int_G (P_\mu^{(R)} f)(gh)\lambda(dg)\rho(dh) \\
&= \int_G (P_\lambda^{(L)} P_\mu^{(R)} f)(h)\rho(dh) \\
&= \int_G (P_\mu^{(R)} P_\lambda^{(L)} f)(h)\rho(dh) \\
&= (P_\mu^* \rho)(P_\lambda^{(L)} f) \\
&\leq \rho(P_\lambda^{(L)} f) = (\lambda * \rho)(f).
\end{aligned}
$$

2. This follows from the easily verified fact that $P_\mu(f \wedge c) \leq P_\mu f \wedge c$. □

Next we establish a connection between properties of convolution operators and existence and regularity of densities as discussed in Sect. 4.5. This result will be useful for us in the next chapter. Readers requiring background on Hilbert-Schmidt operators are referred to Appendix A.6. This result was obtained in Applebaum [9].

Theorem 4.7.1 *Let G be a compact Lie group and $\mu \in \mathcal{P}(G)$. The operator P_μ acting in $L^2(G)$ is Hilbert-Schmidt if and only if μ has a square-integrable density.*

Proof For sufficiency, assume that μ has density $f_\mu \in L^2(G, \mathbb{R})$. Then for all $g \in L^2(G)$, $\sigma \in G$, $(P_\mu g)(\sigma) = \int_G g(\sigma\tau) f_\mu(\tau)d\tau = \int_G g(\tau) f_\mu(\sigma^{-1}\tau)d\tau$. Now define the mapping $k_\mu : G \times G \to \mathbb{R}$ by $k_\mu(\sigma, \tau) := f_\mu(\sigma^{-1}\tau)$. Then $k_\mu \in L^2(G \times G)$ since by left-invariance of (normalised) Haar measure, and Fubini's theorem

$$\int_G \int_G |k_\mu(\sigma, \tau)|^2 d\sigma d\tau = \int_G \int_G |f_\mu(\sigma^{-1}\tau)|^2 d\tau d\sigma = ||f_\mu||_2^2 < \infty,$$

and the result follows by Theorem A.6.4 in Appendix A.6. For necessity, suppose that P_μ is Hilbert-Schmidt. Then it has a kernel $k_\mu \in L^2(G \times G)$ and

$$(P_\mu f)(\sigma) = \int_G f(\tau)k_\mu(\sigma, \tau)d\tau.$$

In particular, for each $A \in \mathcal{B}(G)$,

$$\mu(A) = P_\mu 1_A(e) = \int_A k_\mu(e, \tau)d\tau.$$

Then for all $g \in C(G, \mathbb{R})$, $\int_G g(\sigma)\mu(d\sigma) = \int_G g(\sigma)k_\mu(e, \sigma)d\sigma$. It then follows by the argument used in the last part of the proof of Theorem 4.5.1 that μ is absolutely continuous with respect to m with density $f_\mu := k_\mu(e, \cdot)$, and we also have $f_\mu \in L^2(G, \mathbb{R})$. □

A linear operator $T : B_b(G) \to B_b(G)$ is said to be *strong Feller* if $\mathrm{Ran}(T) \subseteq C_b(G)$. The next theorem can essentially be found in Hewitt and Ross [91] (see (iv) on p. 298). The probabilistic interpretation was first observed by Hawkes [82] for the case $G = \mathbb{R}^n$, where a more detailed analysis appears to be possible.

Theorem 4.7.2 *If $\mu \in \mathcal{P}(G)$ has a continuous density g_μ with respect to left Haar measure, then the convolution operator P_μ is strong Feller.*

Proof We need only establish continuity. Let $\sigma \in G$ and $(\sigma_n, n \in \mathbb{N})$ be a sequence in G that converges to σ. Then for all $f \in B_b(G), n \in \mathbb{N}$,

$$\begin{aligned}
|P_\mu f(\sigma) - P_\mu f(\sigma_n)| &= \left| \int_G f(\sigma\tau)\mu(d\tau) - \int_G f(\sigma_n\tau)\mu(d\tau) \right| \\
&\leq \int_G |f(\tau)||g_\mu(\sigma^{-1}\tau) - g_\mu(\sigma_n^{-1}\tau)|d\tau \\
&\leq \sup_{\tau\in G} |f(\tau)| \int_G |g_\mu(\sigma^{-1}\tau) - g_\mu(\sigma_n^{-1}\tau)|d\tau \\
&\leq \sup_{\tau\in G} |f(\tau)| \int_G (g_\mu(\sigma^{-1}\tau) + g_\mu(\sigma_n^{-1}\tau))d\tau \\
&\leq 2 \sup_{\tau\in G} |f(\tau)|.
\end{aligned}$$

The result then follows by dominated convergence and the continuity of g_μ. □

Finally we establish a useful spectral property for the case where μ is a central measure and G is compact.

Theorem 4.7.3 *If G is compact and $\mu \in \mathcal{P}_c(G)$, then $\{\pi_{ij}, 1 \leq i, j \leq d_\pi, \pi \in \widehat{G}\}$ is a complete set of eigenfunctions for P_μ acting in $L^2(G)$. Moreover, we have*

$$P_\mu \pi_{ij} = \overline{c_\pi} \pi_{ij}$$

for all $1 \leq i, j \leq d_\pi, \pi \in \widehat{G}$, where $\widehat{\mu}(\pi) = c_\pi I_\pi$ (c.f. Corollary 4.2.2).

Proof For all $\sigma \in G$, using Theorem 4.2.1 (4) we have

$$\begin{aligned}
P_\mu \pi_{ij}(\sigma) &= \int_G \pi_{ij}(\sigma\tau)\mu(d\tau) \\
&= \sum_{k=1}^{d_\pi} \pi_{ik}(\sigma) \int_G \pi_{kj}(\tau)\mu(d\tau)
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{k=1}^{d_\pi} \pi_{ik}(\sigma)\widehat{\widetilde{\mu}(\pi)}_{kj} \\
&= \sum_{k=1}^{d_\pi} \pi_{ik}(\sigma)\widehat{\mu(\pi)}^{*}_{kj} \\
&= \sum_{k=1}^{d_\pi} \pi_{ik}\overline{c_\pi}\delta_{kj} \\
&= \overline{c_\pi}\pi_{ij}(\sigma),
\end{aligned}
$$

and the result follows. □

4.8 Recurrence

If $\mu \in \mathcal{P}(G)$ we define $\mu^{*(0)} = \delta_e$. Throughout this section we will assume that $\mu \in \mathcal{P}(G)$ is regular and also *full*, i.e. the closed subgroup of G that is generated by $\mathrm{supp}(\mu)$ is G. We define the *potential measure* V_μ of μ by the prescription

$$
V_\mu := \sum_{n=0}^{\infty} \mu^{*(n)},
$$

so that $V_\mu(f) = \sum_{n=0}^{\infty} \mu(P_\mu^n f)$ for each $f \in C_c(G)_+$. If V_μ is regular, then $V_\mu(f) < \infty$; otherwise it may take the value ∞.

We say that μ is *transient* if $V_\mu(A) < \infty$ for all open relatively compact subsets of G and *recurrent* if $V_\mu(A) = \infty$ for all non-empty open subsets of G. We say that the group G is *recurrent* if $\mathcal{P}(G)$ contains at least one full recurrent measure.

Theorem 4.8.1 (Recurrence-Transience Dichotomy) *Every full $\mu \in \mathcal{P}(G)$ is either recurrent or transient.*

We omit the proof, which can be obtained by combining the results of Theorem 22 (pp. 19–20) and Theorem 26 (pp. 23–24) in Guivarc'h et al. [75].

From the random-walk perspective, recurrence is equivalent to the requirement that for all $g \in G$, $P\left(\lim_{n\to\infty}(S_n \in V_g)|S(0) = e\right) = 1$ for every neighbourhood V_g of g. From the point of view of this monograph a key result is the following:

Proposition 4.8.1 *1. Any full probability measure on a compact group is recurrent.*
2. Every compact group G is recurrent.

Proof 1. Let $\mu \in \mathcal{P}(G)$ be full. By the recurrence-transience dichotomy (Theorem 4.8.1), if $V_\mu(A) = \infty$ for some open relatively compact subset A of G, then G cannot be transient and so must be recurrent. But we may take $A = G$ and then $V_\mu(G) = \sum_{n=0}^{\infty} \mu^{*(n)}(G) = \infty$.

2. Take μ to be normalised Haar measure on G. □

Lemma 4.8.1 *Let μ be a full measure in $\mathcal{P}(G)$ and f be continuous and μ-superharmonic. If μ is recurrent, then f is μ-harmonic.*

Proof We follow Guivarc'h et al [75] p. 43. Without loss of generality, we suppose that $f - P_\mu f > 0$ and seek a contradiction. By the recurrence assumption,

$$\infty = V_\mu(f - P_\mu f) = \sum_{n=0}^{\infty}(P_\mu^n(f - P_\mu f)),$$

and so

$$\infty = \lim_{n\to\infty}(f - P_\mu^n f) \leq f,$$

giving the required contradiction. □

The main result of this section is the following. Our proof closely mirrors that of Guivarc'h et al [75] Proposition 45, pp. 42–44.

Theorem 4.8.2 *Let $\mu \in \mathcal{P}(G)$ be full. The following are equivalent.*

(i) μ is recurrent.
(ii) Every μ-superharmonic continuous function on G is constant.
(iii) Every μ-superharmonic measure is a right Haar measure.
(iv) Every μ-superharmonic function on G is constant m_R-almost everywhere.

Proof (i) $\Rightarrow$ (ii). Let f be a continuous μ-superharmonic function on G. By Lemma 4.8.1 it is μ-harmonic. Assume that f is bounded and suppose that it attains it maximum value, so there exists $g_0 \in G$ such that $f(g_0) = \sup_{g\in G} f(g)$. Then we have

$$f(g_0) = (P_\mu f)(g_0) = \int_G f(g_0 h)\mu(dh),$$

and so

$$\int_G (f(g_0) - f(g_0 h))\mu(dh) = 0.$$

It follows that $f(g_0) = f(g_0 h)$ for all $h \in g_0^{-1}\text{supp}(\mu)$. By repeatedly using the fact that f is harmonic we deduce that $f(g_0) = f(g_0 h_1 \ldots h_n)$ for all $h_1, \ldots, h_n, n \in \mathbb{N}$ such that $g_0 h_1 \ldots h_n \in \text{supp}(\mu)$. Repeatedly applying Proposition 4.7.4, wherein λ is taken to be $\widetilde{\mu}$, allows us to repeat the previous argument with any $h_i, i = 1, \ldots, n$ replaced by h_i^{-1}. Finally, using the continuity of f and the fact that μ is full, we deduce that f is constant. To extend this result to the general case, observe that if $f \neq 0$, we can replace f by $f \wedge c$ where $c > 0$ and appeal to Proposition 4.7.4(2).

(ii) $\Rightarrow$ (i). We assume that μ is transient and seek a contradiction. Let $f \in C_c(G)_+$ be non-trivial and define $F_f = f * V_\mu$, so that for all $g \in G$, $F_f(g) =$

$\sum_{n=0}^{\infty} f(gh)\mu^{*(n)}(dh)$. By a standard use of dominated convergence, we can see that F_f is continuous. Then it is easily verified that

$$P_\mu F_f = F_f - f \leq F_f,$$

and so F_f is both continuous and μ-superharmonic. Hence it is constant, and so $f = P_\mu F_f - F_f = 0$, giving the desired contradiction.

(ii) $\Rightarrow$ (iii). Assume that $\rho \in \mathcal{M}(G)$ is $\widetilde{\mu}$-superharmonic. Choose an arbitrary $\psi \in C_c(G)$ and let α_ψ be the measure of compact support that has Radon-Nikodym derivative $\dfrac{d\alpha_\psi}{dm_R} = \psi$ with respect to right Haar measure. By Proposition 4.7.4(1) $\alpha_\psi * \rho$ is $\widetilde{\mu}$-superharmonic, and it is easily verified that this measure is absolutely continuous with respect to right Haar measure, and has Radon-Nikodym derivative F_ψ where $F_\psi(g) = \int_G \psi(gh^{-1})\rho(dh)$ for all $g \in G$. The mapping F_ψ is clearly non-negative, continuous and it is μ-superharmonic. Then it is constant by (ii). So $F_\psi(g) = F_\psi(e)$ for all $g \in G$. Then for all $\psi \in C_c(G)$, $g \in G$,

$$\int_G \psi(gh^{-1})\rho(dh) = \int_G \psi(h^{-1})\rho(dh).$$

It follows that $\widetilde{\rho}$ is a left Haar measure and so ρ is a right Haar measure.

(iii) $\Rightarrow$ (iv). Let $f \in B_b(G)_+$ and consider the measure $\rho_f \in \mathcal{M}(G)$ for which $\dfrac{d\rho_f}{dm_R} = f$. If f is μ-superharmonic, then ρ_f is $\widetilde{\mu}$-superharmonic, and so ρ_f is a right Haar measure. It follows that f is constant almost everywhere with respect to m_R. If f is not bounded, we can again replace f by $f \wedge c$ where $c > 0$ and use Proposition 4.7.4(2).

(iv) $\Rightarrow$ (ii) is obvious. $\square$

Let M be a Hausdorff topological space and let $\mathcal{P}(M)$ be the space of all regular Borel probability measures on M. A continuous mapping $\alpha : G \times M \to M$ for which $\alpha(e, x) = x$ for all $x \in M$ is called an *action* of G on M. An action is *transitive* if for all $g_1, g_2 \in G$, $x \in M$, $\alpha(g_1, \alpha(g_2, x)) = \alpha(g_1 g_2, x)$. A locally compact Hausdorff group G is said to be *amenable in action*[12] if for every transitive action α on every compact space M, there exists a regular Borel probability measure μ_M on M such that for all $f \in C(M)$, $g \in G$

$$\int_M f(\alpha(g, x))\mu_M(dx) = \int_G f(x)\mu_M(dx).$$

[12] This term should not be confused with the notion of an *amenable group*, which is a group that possesses an invariant mean in the sense of the existence of a positive linear functional on $L^\infty(G)$ that is invariant with respect to left, or right translations. Such groups play an important role in ergodic theory, see e.g. Ornstein and Weiss [154].

The proof of the following theorem is based on that in Guivarc'h et al. [75] (pp. 45–47). We will need the notion of the convolution of $\mu \in \mathcal{P}(G)$ and $\lambda \in \mathcal{P}(M)$ relative to a given transitive action α. This is the measure $\mu *_\alpha \lambda \in \mathcal{P}(M)$ such that for all $A \in \mathcal{B}(M)$:

$$(\mu *_\alpha \lambda)(A) := \int_M \int_G 1_A(\alpha(g, x))\mu(dg)\lambda(dx).$$

Theorem 4.8.3 *Every recurrent group is amenable in action.*

Proof Let μ be a full recurrent probability measure on G and $\nu \in \mathcal{P}(M)$ be arbitrary. For each $n \in \mathbb{N}$ define $\mu_n := \frac{1}{n}\sum_{k=0}^{n-1} \mu^{*(k)}$. Then $(\mu_n *_\alpha \nu, n \in \mathbb{N})$ is weakly relatively compact and so has a subsequence $(\mu_{n_k} *_\alpha \nu, k \in \mathbb{N})$ that converges weakly to some $\lambda \in \mathcal{P}(M)$. If we can prove that λ is invariant, then we are done. We first show that $\mu *_\alpha \lambda = \lambda$. Indeed, we have for all $f \in C(M)$, using the transitivity of α,

$$\begin{aligned}
(\mu *_\alpha \lambda)(f) &= \int_M \int_G f(\alpha(g, x))\mu(dg)\lambda(dx) \\
&= \lim_{k\to\infty} \int_M \int_G \int_G f(\alpha(gh, x))\mu(dg)\mu_k(dh)\nu(dx) \\
&= \lim_{k\to\infty} \int_M f(x)[(\mu * \mu_k) *_\alpha \nu](dx) \\
&= \lim_{k\to\infty} \int_M f(x)(\mu_k *_\alpha \nu)(dx) = \lambda(f).
\end{aligned}$$

To see that λ is indeed invariant, let $f \in C(M)_+$ and define $\Phi_f \in C(G)_+$ by $\Phi_f(g) = \int_M f(\alpha(g, x))\lambda(dx)$ for $g \in G$. Then Φ_f is μ-harmonic, for by Fubini's theorem

$$\begin{aligned}
P_\mu \Phi_f(g) &= \int_M \int_G f(\alpha(gh, x))\mu(dh)\lambda(dx) \\
&= \int_M f(\alpha(g, x))(\mu *_\alpha \lambda)(dx) \\
&= \int_M f(\alpha(g, x))\lambda(dx) = \Phi_f(g).
\end{aligned}$$

So by Theorem 4.8.2, Φ_f is constant, and hence for all $g \in G$,

$$\int_M f(\alpha(g,x))\lambda(dx) = \Phi_f(g) = \Phi_f(e) = \int_M f(\alpha(e,x))\lambda(dx) = \int_M f(x)\lambda(dx),$$

and the result follows. □

For example, let H be a closed subgroup of a compact group G and $M = G/H$ be the (compact) homogeneous space of left cosets. Define the *natural action* of G on M by $\alpha(g, g'H) = gg'H$ for all $g, g' \in G$. This is clearly continuous and transitive. G is recurrent by Proposition 4.8.1, and so by Theorem 4.8.3 we can assert the existence of $\mu_M \in \mathcal{P}(M)$ so that for all $g \in G$

$$\int_M f(gx)\mu_M(dx) = \int_M f(x)\mu_M(dx).$$

In particular take $G = SO(n)$ and $H = SO(n-1)$, so that $M = S^{n-1}$. In this case μ_M is the normalised surface measure σ_{n-1}, for which we have the recursive formula

$$\int_{S^{n-1}} f(x)\sigma_{n-1}(dx)$$
$$= \frac{\Gamma\left(\frac{n}{2}\right)}{\sqrt{\pi}\Gamma\left(\frac{n-1}{2}\right)} \int_0^\pi \left(\int_{S^{n-2}} f(\sin(\theta y) + \cos(\theta)e_n)\sigma_{n-2}(dy) \right) \sin^{n-2}(\theta)d\theta,$$

for all $f \in C(S^{n-1})$ where e_n is the "north pole" in S^{n-1} (see Faraut [63], pp. 186–190 for details).

Some detailed results on recurrent random walks on non-compact groups (at least in the abelian case) may be found in Chap. 9 of Revuz [167].

4.9 Notes and Further Reading

The interaction between probability and group theory covers a huge area and it is difficult to do this justice in such a short space. In particular this includes probability on discrete groups (which is not really the topic of this monograph) and has considerable overlap with "stochastic differential geometry", as random motion on a Lie group can be regarded as a special case of that on a more general manifold.

In the 1960s a considerable literature began to evolve on probability theory in such general mathematical structures as groups on the one hand and Banach spaces on the other. These two directions have now diverged considerably, but in 1963 Grenander [73] was able to justify including both themes within a single volume. From the continuous group point of view, he introduced the Fourier transform and gave some attention to limit theorems. He traces the historic roots of the subject back to work by

Perrin [159] in 1928 on Brownian motion in the rotation group. Hannan's survey paper [78] from 1965 is also of historical interest. He develops applications to second-order stationary processes, experimental design and ANOVA. Parthasarathy's book [155], which appeared in 1967, discusses probability on metric groups and gives an account of infinite divisibility, the Lévy-Khintchine formula and the central limit theorem on locally compact abelian groups. Ten years later, Heyer's highly influential treatise [95] appeared which gave a comprehensive and detailed state of the art account of probability on (general) locally compact groups. This monograph is now a classic and after 35 years is still a highly valuable resource for those doing research in this area. Highlights are the treatment of Hunt's classification of the generators of convolution semigroups and the central limit theorem. We will investigate both of these topics in the next chapter. Ten years later, Diaconis [56] published his beautiful lecture notes that demonstrate the fruitfulness of group actions in a variety of contexts within both probability and statistics, from card shuffling to ANOVA and spectral analysis of time series.

In more recent years there have been a number of books and monographs on more specific topics concerned with probability on groups. For example Hazod and Siebert [85] study stable laws on locally compact groups (where these make sense), Neuenschwander [153] investigates limit theorems and Brownian motion on the Heisenberg group, and Liao's monograph [132] is devoted to Lie group-valued Lévy processes. The treatise of Guivarc'h et al. [75] is a comprehensive account of random walks on groups. For a more recent survey, see Breuillard [35]. Random walks on the rotation groups $SO(n)$ were given an extensive treatment by Rosenthal [170]. For a study of random walks on spheres, see Bingham [29]. Central measures were introduced in this chapter and will feature prominently in the next one. These have been investigated by a number of analysts (see e.g. Ragozin [163] and Hare [79]) and probabilists (see e.g. Siebert [184]).

We can regard normalised Haar measure on $U(n)$ as the uniform distribution therein, and choosing unitary matrices according to this law plays an important role in *random matrix theory*; see e.g. Keating and Snaith [115] for intriguing connections to the Riemann hypothesis, and Diaconis and Shahshahani [57] for computations of the eigenvalues of random matrices in connection with a continuous generalisation of the classical matching problem. For a monograph treatment of this topic, see Anderson et al. [3].

It is worth pointing out that there is an interesting class of locally compact groups called *Moore groups*, whose defining property is that all of their irreducible representations are finite dimensional. So all compact groups are Moore groups and many probabilistic results that hold for Moore groups are automatically applicable to compact groups; see e.g. Sects. 1.3 and 1.4 of Heyer [95].

http://www.springer.com/978-3-319-07841-0

Probability on Compact Lie Groups
Applebaum, D.
2014, XXVI, 217 p., Hardcover
ISBN: 978-3-319-07841-0