ILLINOIS JOURNAL OF MATHEMATICS
Volume 38, Number 4, Winter 1994

ERGODIC THEOREMS FOR CONVOLUTIONS OF A MEASURE ON A GROUP

ROGER JONES[1], JOSEPH ROSENBLATT[1] AND ARKADY TEMPELMAN

Introduction

Let G be a Hausdorff locally compact group (called a group here) and let μ be a probability measure in $M(G)$, the finite regular Borel measures on G. By $\|\mu\|_1$, we will denote the total variation norm of $\mu \in M(G)$. Suppose that (X, β, m) is a measure space with m being a σ-finite positive measure. Let T be a representation of G as invertible measure-preserving transformations of (X, β, m). Then there is an operator on $L_2(X, \beta, m)$ associated with μ, denoted by T_μ, which integrates T_g, $g \in G$, with respect to μ. This operator can be defined weakly by

$$\langle T_\mu f_1, f_2 \rangle = \int\int f_1(T_{g^{-1}}x)\overline{f_2(x)}\, dm(x)\, d\mu(g)$$

for all $f_1, f_2 \in L_2(X, \beta, m)$. In the books by Tempelman [34], [35] and in several recent articles (see Bellow, Jones, and Rosenblatt [3], [4], [5], Derriennic and Lin [9], and Rosenblatt [28, 29]) in the case of probability measures μ, the norm and almost everywhere behavior of the iterates of T_μ on $L_p(X, \beta, m)$ have been studied with some success. In this article, these various results are extended to general locally compact groups, including a specific discussion of the influence that the spectral behavior of μ and T_μ have on the conclusions. Various positive results about norm and a.e. convergence of the iterates of T_μ will be obtained, and counterexamples will be discussed which illustrate the limitations on the theorems and the techniques that are used.

1. Direct integral formulas

The first issue is to clarify the definition of $T_\mu = \int T_g\, d\mu(g)$. If $d\mu = \phi\, d\lambda_G$, where λ_G is a right-invariant Haar measure on G, and $\phi \in L_1(G, \lambda_G)$, then this operator is the standard integration of the representation T as unitary

Received July 22, 1992.
1991 Mathematics Subject Classification. Primary 28; Secondary 43.
[1]Partially supported by NSF grants.

operators on $L_2(X, \beta, m)$ defined by $T_g f(x) = f(T_{g^{-1}}x)$ for all $f \in L_2(X, \beta, m)$. See Hewitt and Ross [17]. To define T_μ more generally requires only a use of the Fubini theorem. For this reason, we assume that G is σ-compact and that T is jointly measurable in the sense that the mapping $G \times X \to X$ by $(g, x) \to T_g x$, $g \in G$, $x \in X$, is measurable with respect to the product σ-algebra $\beta_G \times \beta$ in $G \times X$, where β_G denotes the Borel sets in G. See Rudin [30] for technical details stated here that arise in the use of Fubini's theorem.

It is going to be necessary at some points in the sequel to know something about the continuity of the representation $g \mapsto T_g$. Generally, if $L_2(X, \beta, m)$ is separable, then the measurability of this representation gives weak and strong continuity. By approximation, it follows that T is continuous on any $L_p(X, \beta, m)$, $1 < p < \infty$. However, T is also continuous on $L_1(X, \beta, m)$ because $L_1(X, \beta, m)$ is separable too. See Moore [24]. It is also the case that the assumption of joint measurability of T implies that T is continuous on each $L_p(X, \beta, m)$, $1 \le p < \infty$. In any case then, there is no harm in assuming the continuity throughout in what follows. See Tempelman [34, 35] for further information on these issues.

Now fix p, $1 \le p \le \infty$, and let q be the index conjugate to p. Let $f_1 \in L_p(X, \beta, m)$ and let $f_2 \in L_q(X, \beta, m)$. Then define

$$\mathscr{G}(f_1, f_2)(g) = \int f_1(T_{g^{-1}}x)\overline{f_2(x)}\, dm(x) \quad \text{for all } g \in G.$$

This function is well-defined and $\|\mathscr{G}(f_1, f_2)\|_\infty \le \|f_1\|_p \|f_2\|_q$. Thus, $\mathscr{G}(f_1, f_2)$ is a bounded Borel measurable function on G. Thus, we can define

$$I(f_1, f_2) = \int \mathscr{G}(f_1, f_2)(g)\, d\mu(g).$$

This gives a sesquilinear form I such that

$$|I(f_1, f_2)| \le \|f_1\|_p \|f_2\|_q \|\mu\|_1.$$

Also, this shows that the function $(g, x) \mapsto f_1(T_{g^{-1}}x)\overline{f_2(x)}$ is in $L_1(G \times X, \beta_G \times \beta, \mu \times m)$ and for a.e. $x \in X$, $g \mapsto f_1(T_{g^{-1}}x)\bar{f}_2(x)$ is in $L_1(G, \beta_G, \mu)$. Also, the function

$$X(f_1, f_2)(x) = \int f_1(T_{g^{-1}}x)\overline{f_2(x)}\, d\mu(g)$$

is in $L_1(X, \beta, m)$ with $I(f_1, f_2) = \int X(f_1, f_2)\, dm$. Hence, I can be computed by either iterated integral.

This argument shows that for all $f_1 \in L_p(X, \beta, m)$, there exists a unique element, denoted $T_\mu f_1$, in $L_p(X, \beta, m)$ such that $I(f_1, f_2) = \int T_\mu f_1 \bar{f}_2 \, dm$. Also, for a.e. x,

$$T_\mu f_1(x) = \int f_1(T_{g^{-1}} x) \, d\mu(g).$$

It is clear that T_μ is a bounded linear operator on $L_p(X, \beta, m)$ with $\|T_\mu\| \le \|\mu\|_1$. The operator T_μ defined in this manner is what is meant by $\int T_g \, d\mu(g)$.

In a manner similar to the above, if T is a continuous representation of G as unitary operators on a Hilbert space H, then for any $\mu \in M(G)$, there is a continuous operator T_μ defined on H weakly by $\langle T_\mu f_1, f_2 \rangle = \int \langle T_g f_1, f_2 \rangle \, d\mu(g)$ for $f_1, f_2 \in H$. Then $\|T_\mu\| \le \|\mu\|_1$ too. If $H = L_2(X, \beta, m)$, the previous use of Fubini's theorem extends this to giving an interpretation of $T_\mu f_1$ as an integral of the form $T_\mu f_1(x) = \int f_1(T_{g^{-1}} x) \, d\mu(g)$ for a.e. $x \in X$. If T is an irreducible representation λ of G on H, then T_μ will be denoted in the sequel by T_μ^λ to emphasize the dependence on λ, even though λ_μ would be more strictly consistent with the previous notation. The reason for this notation is more apparent in stating Corollary 1.6.

Now if $\mu \ge 0$, then $T_\mu \ge 0$ and if μ is a probability measure then T_μ is a positive contraction simultaneously on all L_p spaces. Moreover, if μ is the Dirac mass δ_g at $g \in G$, then $T_\mu = T_g$. The mapping $\mu \mapsto T_\mu$ from $M(G)$ to the bounded operators on $L_p(X, \beta, m)$ is a well-defined linear mapping. It is also a Banach algebra homomorphism commuting with the adjoint operator. The routine proof of this fact which is stated precisely in 1.1 will be omitted.

1.1. PROPOSITION. *If $\mu, \nu \in M(G)$, then $T_{\mu * \nu} = T_\mu \circ T_\nu$. If μ^* is defined by $\mu^*(E) = \mu(E^{-1})$ for all $E \in \beta_G$, then $(T_\mu)^* = T_{\mu^*}$.*

1.2. *Remark.* Just as above, if T is a continuous representation of G as unitary operators on a Hilbert space H, then $\mu \mapsto T_\mu$ is a Banach $*$-algebra homomorphism of $M(G)$ into the bounded operators on H. We will need this observation later in discussing the direct integral decomposition of T_μ.

For applications in ergodic theory, one important task is to be able to compute the spectrum of T_μ on $L_2(X, \beta, m)$. To carry this out via direct integrals, we need some separability hypothesis on G. By $\hat{G}$ is meant the unitary dual of G in the Fell topology. See Dixmier [10] and Fell [11], [12], [13]. The separability of $\hat{G}$ will be needed in Theorem 1.5 and the use of Choquet's Theorem in Corollary 1.7.

1.3. PROPOSITION. *If G is a σ-compact group, then G has a metric topology if and only $\hat{G}$ is separable* (*i.e., has a countable dense subset.*)

Proof. If G is σ-compact and metric then $L_2(G, \beta_G, \lambda_G)$ is separable. Therefore, the convex set $P \subset L_\infty(G, \beta_G, \lambda_G)$ of regular positive linear functionals on $L_1(G, \beta_G, \lambda_G)$ with norm no more than 1 is separable in the w^*-topology. See Hewitt and Ross [17] for a discussion of P. It is immediate that $\hat{G}$ is then separable. Conversely, if $\hat{G}$ is separable, let $\Gamma \subset \hat{G}$ be a countable dense subset. Because G is σ-compact and each $\gamma \in \Gamma$ is irreducible, there is a countable dense subset $V_\gamma \subset H_\gamma$ in the underlying Hilbert space H_γ of γ. For each finite set $A \subset \Gamma$, finite sets $B_\gamma \subset V_\gamma$, $\gamma \in A$, and $k \geq 1$, let

$$N\big(A, \{B_\gamma\}, k\big) = \big\{g \in G: |\langle \gamma(g)v, v\rangle - \langle v, v\rangle| < 1/k \text{ for all } \gamma \in A \text{ and } v \in B_\gamma\big\}.$$

Then the sets $N(A, \{B_\gamma\}, k)$ are open neighborhoods of $e \in G$, and there are countably many of them. If we show they form a neighborhood basis at $e \in G$, then G is first axiom and metric. First, any g in all $N(A, \{B_\gamma\}, k)$ has $\langle \gamma(g)v, v\rangle = \langle v, v\rangle$ for all $\gamma \in \Gamma$ and $v \in V_\gamma$. But then $\gamma(g) = \mathrm{Id}$ for $\gamma \in \Gamma$. However, the density of Γ in $\hat{G}$ says that if $\rho \in \hat{G}$, and $w \in H_\rho$, there exists $\gamma_\alpha \in \Gamma$, $w_\alpha \in H_{\gamma_\alpha}$, such that $\lim_\alpha \langle \gamma_\alpha(g)w_\alpha, w_\alpha\rangle = \langle \rho(g)w, w\rangle$ uniformly on compacta in G. Hence, $\langle \rho(g)w, w\rangle = \langle w, w\rangle$ for any $w \in H_\rho$ and so $\rho(g) = \mathrm{Id}$ for all $\rho \in \hat{G}$. That is, if g is all $N(A, \{B_\gamma\}, k)$ then $g = e$. But now take any compact neighborhood K of e in G and any open set $U, e \in U \subset K$. If no $N(A, \{B_\gamma\}, k) \subset U$, then by the finite-intersection property there is some $g \in K \setminus U$ which is in all $N(A, \{B_\gamma\}, k)$. However then $g \in e$ and $g \in U$, a contradiction which shows the sets $N(A, \{B_\gamma\}, k)$ are a basis of neighborhoods of e in G. Thus, G is first axiom and a metric space. □

1.4. *Remark.* It is also clear that G is a σ-compact metric group if and only if $L_2(G, \beta_G, \lambda_G)$ is separable. Some authors say G is separable in this case. This can be confusing. For example, the compact abelian group $\prod_T T, T = \{z \in \mathbf{C}: |z| = 1\}$, is separable (in that it has a countable dense subset), but it is not a metric group and $\hat{G}$ is not separable. Also, generally if $\hat{G}$ is separable, then G is a metric group, but the converse is not the case without some extra hypothesis, like σ-compactness.

The spectral decomposition theorem that we need is Theorem 1.5. It requires the strong separability hypothesis that $L_2(G, \beta_G, \lambda_G)$ has a countable dense set, which Propostion 1.3 and Remark 1.4 should clarify. There are several versions of Theorem 1.5, some global as mentioned in Tempelman [34], [35]. However, we need only the version given here. See Mackey [22] and Dixmier [10] for good discussions of this existence theorem. These sources also explain the measurability concepts used here.

1.5. THEOREM. *Given a σ-compact metric group, and a continuous unitary representation T of G, there is a measure space $(\Lambda, \mathscr{D})$ consisting of continuous irreducible representations over which T is unitarily equivalent to a direct integral $\int_\Lambda \oplus T^\lambda \, dF(\lambda)$ with respect to some σ-finite positive measure F on $(\Lambda, \mathscr{D})$ and some measurable $\lambda \mapsto T^\lambda$ from Λ to $\hat{G}$.*

There are many technical issues associated with such direct integral decompositions. For example, the measure space $(\Lambda, \mathscr{D})$ can be chosen without dependence on T, so that only F remains to be chosen given T. This is not important in our applications. We only really need the fact that T_μ also decomposes as a direct integral via the above. Also, no uniqueness of this decomposition is needed, which is fortunate since only certain groups have such a property. See Mackey [22] and Tempelman [34], [35] for the following. For additional facts about direct integrals, see also Dixmier [10].

1.6. COROLLARY. *If G is a σ-compact metric group, and $\mu \in M(G)$, then with respect to a direct integral decomposition of T, we have $T_\mu = \int_\Lambda \oplus T_\mu^\lambda \, dF(\lambda)$.*

In particular, for G abelian, the representation space H_λ of each T^λ, $\lambda \in \Lambda$, is one-dimensional and $T_\mu = \int_\Lambda \oplus \hat{\mu}(\lambda)\mathrm{Id} \, dF(\lambda)$.

In any of these direct integral decompositions of μ, if $\mu \geq 0$, then F is positive. However, the formula is also useful when μ is some polynomial in a probability measure because of the following. Suppose T is written as a direct integral as above and $f \in L_2(X, \beta, m)$. Then f is $\int \oplus f_\lambda \, dF(\lambda)$ for some $f_\lambda \in H_\lambda$, $\lambda \in \Lambda$, and $\|f\|_2^2 = \int_\Lambda \|f_\lambda\|_2^2 \, dF(\lambda)$. It follows that

$$\|T_\mu f\|_2^2 = \int_\Lambda \left\| T_\mu^\lambda f_\lambda \right\|_2^2 dF(\lambda)$$

and so

$$\|T_\mu\| \leq \sup_{\lambda \in \Lambda} \left\| T_\mu^\lambda \right\|.$$

This is a key fact in the following and so we state it formally here and give another proof of it, one which avoids the language of direct integrals.

1.7. COROLLARY. *If G is a σ-compact metric group, and $\mu \in M(G)$, then $\|T_\mu\| \leq \sup_{\lambda \in \Lambda} \|T_\mu^\lambda\|$.*

Proof. Let $f \in L_2(X, \beta, m)$, $\|f\|_2 \leq 1$. Then $p_0(g) = \langle T_g f, f \rangle$ is a continuous positive definite function on G. As in Proposition 1.3, P is separable and so we can apply Choquet's Theorem to see there is a probability measure

σ on the extreme points EP of the set of continuous positive definite functions P on G such that $p_0 = \int_{EP} p \, d\sigma(p)$. That is, for $f \in L_1(G, \beta_G, \lambda_G)$,

$$\int f(g) p_0(g) \, d\lambda_G(g) = \int_{EP} \int f(g) p(g) \, d\lambda_G(g) \, d\sigma(p).$$

It follows that $p_0(g) = \int_{EP} p(g) \, d\sigma(p)$ for $g \in G$ by Fubini's theorem. So

$$\begin{aligned} \langle T_\mu f, T_\mu f \rangle &= \langle T_{\mu^* * \mu} f, f \rangle \\ &= \int_G \langle T_g f, f \rangle d(\mu^* * \mu)(g) \\ &= \int_G \int_{EP} p(g) \, d\sigma(p) d(\mu^* * \mu)(g). \end{aligned}$$

Now, each $p(g) = \langle T_g^\gamma \nu_\gamma, \nu_\gamma \rangle$ for some continuous irreducible representation $\gamma = \gamma(p)$ of G on a Hilbert space H_γ with $\nu_\gamma \in H_\gamma$, $\|\nu_\gamma\|_{H_\gamma} \leq 1$. Hence,

$$\begin{aligned} \int_G p(g) d(\mu^* * \mu)(g) &= \int_G \langle T_g^\gamma \nu_\gamma, \nu_\gamma \rangle d(\mu^* * \mu)(g) \\ &= \langle T_{\mu^* * \mu}^\gamma \nu_\gamma, \nu_\gamma \rangle \\ &= \langle T_\mu^\gamma \nu_\gamma, T_\mu^\gamma \nu_\gamma \rangle \\ &= \|T_\mu^\gamma \nu_\gamma\|^2 \leq \|T_\mu^\gamma\|^2. \end{aligned}$$

Thus,

$$\begin{aligned} \langle T_\mu f, T_\mu f \rangle &= \int_{EP} \int_G p(g) d(\mu^* * \mu)(g) \, d\sigma(p) \\ &\leq \int_{EP} \sup_{\gamma \in \hat{G}} \|T_\mu^\gamma\|^2 \, d\sigma(p) \\ &= \sup_{\gamma \in \hat{G}} \|T_\mu^\gamma\|^2. \end{aligned}$$

So $\|T_\mu\|^2 \leq \sup_{\gamma \in \hat{G}} \|T_\mu^\gamma\|^2$. $\square$

1.8. *Remark.* Notice that $\|T_\mu\| = \sup_{\gamma \in \hat{G}} \|T_\mu^\gamma\|$ for every μ if the representation T weakly contains every irreducible representation $\gamma \in \hat{G}$.

Another aspect of the direct integral representation is that it gives better criteria for the invertibility of T_μ than one can get using just criteria for the invertibility of μ in $M(G)$. For example:

1.9. PROPOSITION. *If G is a σ-compact, metric abelian group, and* $\inf\{|\hat{\mu}(\lambda)|: \lambda \in \hat{G}\} > 0$, *then T_μ is invertible. Hence the spectrum of T_μ is a subset of* $\mathrm{cl}_{\mathbf{C}}\{\hat{\mu}(\lambda): \lambda \in \hat{G}\}$.

Proof. Here $T_\mu = \int_\Lambda \oplus \hat{\mu}(\lambda)\,\mathrm{Id}\, dF(\lambda)$. Hence, if $\inf\{|\hat{\mu}(\lambda)|: \lambda \in \hat{G}\} > 0$, then

$$T_\mu^{-1} = \int_\Lambda \oplus \hat{\mu}(\lambda)^{-1}\,\mathrm{Id}\, dF(\lambda)$$

is a well-defined bounded inverse for T_μ. If $\gamma \notin \mathrm{cl}_{\mathbf{C}}\{\hat{\mu}(\lambda): \lambda \in \hat{G}\}$, then $\gamma\,\mathrm{Id} - T_\mu = T_{\gamma\delta_e - \mu}$ has $|\gamma\hat{\delta}_e - \hat{\mu}|$ bounded away from zero. So $\gamma\,\mathrm{Id} - T_\mu$ is invertible. That is,

$$\mathbf{C} \setminus \mathrm{cl}_{\mathbf{C}}\{\hat{\mu}(\lambda): \lambda \in \hat{G}\} \subset \mathbf{C} \setminus sp(T_\mu). \qquad \square$$

1.10. *Remark.* (a) If G is a discrete abelian group, or, more generally, if G is abelian and μ has no singular part, then the Wiener-Pitt Theorem says that the criteria above gives the invertibility of μ in $M(G)$. Hence, by Proposition 1.1, T_μ is invertible too. However, the Wiener-Pitt phenomenon (see Rudin [31] and Williamson [37]) shows that μ can fail to be invertible while T_μ is invertible.

(b) If G is not abelian, then the same principle as above applies if each T_μ^λ is normal because if γ is a distance ε from $sp(R)$ for a normal operator R, then $(\gamma\,\mathrm{Id} - R)^{-1}$ exists and

$$\|(\gamma\,\mathrm{Id} - R)^{-1}\| \leq 1/\varepsilon.$$

However, failing normality, it is possible for γ to be separated by ε from each $sp(T_\mu^\lambda)$ and each $(\gamma\,\mathrm{Id} - T_\mu^\lambda)^{-1}$ to exist without there being a bound on $\|(\gamma\,\mathrm{Id} - T_\mu^\lambda)^{-1}\|$ independent of λ, which would allow $\gamma\,\mathrm{Id} - T_\mu$ to exist too. This is certainly abstractly the case, but good examples of this possibility occurring in this explicit a context are needed.

2. Ergodic theorems for abelian groups

Various criteria for the norm and/or almost everywhere convergence of the powers of T_μ when $G = Z$ have been given which have been successful because of the ability to compute the spectral behavior of T_μ via $\hat{\mu}$. Some of these results can be generalized.

2.1. Definition. For a probability measure $\mu \in M(G)$, we say μ is *adapted* if the support of μ generates a dense subgroup of G, and we say μ is *strictly aperiodic* if the support of μ is not contained in a proper closed left coset of G.

See Rosenblatt [27] for the background and the information below. See also Glasner [16].

2.2. Proposition. *Let G be an abelian group.*
1) *μ is adapted if and only if $\lambda \in \hat{G}$ and $\hat{\mu}(\lambda) = 1$ implies $\lambda = 1$.*
2) *μ is strictly aperiodic if and only if $\lambda \in \hat{G}$ and $|\hat{\mu}(\lambda)| = 1$ implies $\lambda = 1$.*

When G is a discrete abelian group, or when μ is spread-out (i.e., some μ^n and λ_G are not mutually singular), then μ is strictly aperiodic if and only if μ is adapted and

$$\lim_{n\to\infty} \|\mu^{n+1} - \mu^n\|_1 = 0.$$

2.3. *Example.* (a) Let a, b, c be rationally independent real numbers. Let $\mu = \frac{1}{3}(\delta_a + \delta_b + \delta_c)$. Then $\|\mu^{n+1} - \mu^n\|_1 = 2$ for all $n \geq 1$, but μ is strictly aperiodic.
(b) The example μ in Rudin [31] is a continuous probability measure, singular to Lebesgue measure, such that μ is strictly aperiodic, but $\|\mu^{n+1} - \mu^n\|_1 1 = 2$ for all $n \geq 1$.
Here is one simple general principle for abelian groups. This follows also from the result in Blum-Eisenberg [6]; see Tempelman [34, 35].

2.4. Theorem. *If G is abelian and μ is strictly aperiodic, then for m finite, and all $f \in L_p(X, \beta, m)$, $1 \leq p < \infty$, $\lim_{n\to\infty} T_\mu^n f = P_I f$, the projection of f on the G-invariant functions, with convergence in the L_p-norm.*

Proof. By general principles, it suffices to prove

$$lim_{n\to\infty} \|T_\mu^n(f - T_\mu f)\|_2 = 0$$

for all $f \in L_2$. There is also no harm in identifying Λ as $\hat{G}$ in this case. But if $f = \int \oplus f_\gamma \, dF(\gamma)$, then

$$\left\|T_{\mu^n}(f - T_\mu f)\right\|_2^2 \leq \int \left|\hat{\mu}^{n+1}(\gamma) - \hat{\mu}^n(\gamma)\right|^2 \|f_\gamma\|_2^2 \, dF(\gamma).$$

Here $|\hat{\mu}^n(\gamma)| < 1$ for $\gamma \neq 1$. So as $n \to \infty$, $|\hat{\mu}^{n+1}(\gamma) - \hat{\mu}^n(\gamma)|^2 \|f_\gamma\|_2^2 \to 0$ for all γ. By the Lebesgue Dominated Convergence Theorem,

$$lim_{n \to \infty} \|T_\mu^n(f - T_\mu f)\|_2 = 0. \qquad \square$$

2.5. *Remark.* A similar result is true for m just σ-finite in $L_p(X, \beta, m)$, $1 < p < \infty$. However, it will not be true in $L_1(X, \beta, m)$. Indeed, if T is ergodic, then $I \cap L_1(X, \beta, m) = \{0\}$ if m is not finite. However, for any $f \in L_1(X, \beta, m)$, $f \geq 0$, and any $\mu \geq 0$, $\|T_\mu^n f\|_1 = \|f\|_1 \mu(G)$ for all $n \geq 1$. See Section 4 for further discussion.

By analogy, the cousin to 2.4 is the following:

2.6. THEOREM. *Let T be a normal contraction operator on $L_2(X, \beta, m)$ and suppose that the resolution of the identity E corresponding to T has $E\{z: |z| = 1, z \neq 1\} = 0$. Then $T^n f$ converges in L_2-norm to $E_{\{1\}} f$ for all $f \in L_2(X, \beta, m)$.*

Here is an alternative proof of Theorem 2.4 which does not use the formalism of direct integrals. Assume G acts continuously as unitary operators on a Hilbert space H and that $I = \{f \in H: f \text{ is } G\text{-invariant}\}$. To see Theorem 2.4, we take $H = L_2(X, \beta, m)$ and show that for μ strictly aperiodic and $f \in I^\perp$, $\lim_{n \to \infty} \|T_\mu^n f\| = 0$. But for any $f \in H$, $p(g) = \langle T_g f, f \rangle$ is a continuous positive-definite function on G. So, because G is abelian, the Bochner-Weil theorem says there is a positive finite measure μ_f on $\hat{G}$ with

$$p(g) = \int_{\hat{G}} \overline{\gamma(g)}\, d\mu_f(\gamma) \quad \text{for all} \quad g \in G.$$

The Bochner-Weil theorem can be viewed as a substitute in this case for Choquet's theorem as it was used in the proof of 1.7.

Now generally, the element w in the closed convex hull of $\{T_g f: g \in G\}$ which has smallest norm, is G-invariant. So when $f \in I^\perp$, $w = 0$. That is, there exists a sequence of convex sums $\sum \lambda_i(s) T_{g_i(s)} f$ such that $\lim_{s \to \infty} \|\sum \lambda_i(s) T_{g_i(s)} f\| = 0$. But then,

$$\begin{aligned} 0 &= \lim_{s \to \infty} \left\langle \sum \lambda_i(s) T_{g_i(s)} f, \sum \lambda_i(s) T_{g_i(s)} f \right\rangle \\ &= \lim_{s \to \infty} \sum \lambda_i(s) \lambda_j(s) \left\langle T_{g_j^{-1}(s) g_i(s)} f, f \right\rangle \\ &= \lim_{s \to \infty} \int \sum \lambda_i(s) \lambda_j(s) \overline{\gamma\left(g_j^{-1}(s) g_i(s)\right)}\, d\mu_f(\gamma) \\ &= \lim_{s \to \infty} \int_{\hat{G}} \left| \sum \lambda_i(s) \gamma(g_i(s)) \right|^2 d\mu_f(\gamma). \end{aligned}$$

Since the integrand here is 1 at $\gamma = 1$, and $\mu_f \geq 0$, we see this shows $\mu_f\{1\} = 0$.

But now we can also calculate,

$$\begin{aligned}\langle T_\mu^n f, T_\mu^n f\rangle &= \langle T_{\mu^* * \mu}^n f, f\rangle \\ &= \int \langle T_g f, f\rangle d(\mu^* * \mu)^n(g) \\ &= \int\int \overline{\gamma(g)}\, d\mu_f(\gamma) d(\mu^* * \mu)^n(g) \\ &= \int\int \overline{\gamma}(g) d(\mu^* * \mu)^n(g)\, d\mu_f(\gamma) \\ &= \int \left(\widehat{\mu^* * \mu}\right)^n(\gamma)\, d\mu_f(\gamma) \\ &= \int |\hat{\mu}(\gamma)|^{2n}\, d\mu_f(\gamma).\end{aligned}$$

Since $|\hat{\mu}(\gamma)| < 1$ for $\gamma \neq 1$, and $\mu_f\{1\} = 0$, letting $n \to \infty$ gives $\lim_{n\to\infty} \|T_\mu^n f\|_2^2 = 0$ for any $f \in I^\perp$. That is, $T_\mu^n f \to P_I f$ as $n \to \infty$ for all $f \in H$.

The same technique can be used to prove 1.9. Indeed, with the notation above

$$\begin{aligned}\|T_\mu f\|^2 &= \int_{\hat{G}} |\hat{\mu}(\gamma)|^2\, d\mu_f(\gamma) \\ &\geq \inf_{\gamma \in \hat{G}} |\hat{\mu}(\gamma)|^2 \mu_f(\hat{G}).\end{aligned}$$

But $\|f\|^2 = \langle T_e f, f\rangle = \int \overline{e(\gamma)}\, d\mu_f(\gamma) = \mu_f(\hat{G})$. So

$$\|T_\mu f\| \geq inf_{\gamma \in \hat{G}} |\hat{\mu}(\gamma)|\, \|f\|.$$

Thus, if this infimum is not zero, we know T_μ is 1-1, and, as a simple argument with Cauchy sequences shows, Range(T_μ) is closed. But then also Range$(T_\mu) = H$, since otherwise there is $h \in H$, $h \neq 0$, such that $0 = \langle T_\mu f, h\rangle$ for all $f \in H$. But then $T_{\mu^*} h = 0$ while T_{μ^*} also satisfies

$$\|T_{\mu^*} h\| \geq \inf_{\gamma \in \hat{G}} \left|\widehat{\mu^*}(\gamma)\right| \|h\| > 0,$$

because $\widehat{\mu^*}(\gamma) = \overline{\hat{\mu}(\gamma)}$ for all $\gamma \in \hat{G}$. That is, T_μ is 1-1 and onto, and thus is invertible.

To get results in the abelian case which apply to the issue of pointwise convergence requires more from μ than strict aperiodicity. For example, the reason that the criterion

$$\lim_{n\to\infty} \|\mu^{n+1} - \mu^n\|_1 = 0$$

is important can be seen from the following.

2.7. PROPOSITION. *Suppose G is an abelian group and $\lim_{n\to\infty}\|\mu^{n+1} - \mu^n\|_1 = 0$. Then for all $f \in L_\infty(X, \beta, m)$, $\lim_{n\to\infty}\|T_\mu^n(f - T_\mu f)\|_\infty = 0$.*

Proof. $\|T_\mu^n(f - T_\mu f)\|_\infty = \|T_{\mu^{n+1}-\mu^n} f\|_\infty \leq \|\mu^{n+1} - \mu^n\|_1\|f\|_\infty$. $\square$

2.8. PROPOSITION. *Suppose G is an abelian group and $\lim_{n\to\infty}\|\mu^{n+1} - \mu^n\|_1 = 0$. Assume m is finite. Then on the subspace*

$$S = \{f_1 + f_2 - T_\mu f_2 \colon f_1 \text{ is } G\text{-invariant}, f_1 \in L_\infty(X, \beta, m),$$
$$f_2 \in L_\infty(X, \beta, m)\},$$

which is dense in $L_p(X, \beta, m)$ for any $1 \leq p < \infty$, we have if $f \in S$, $\lim_{n\to\infty} T_{\mu^n} f(x)$ exists a.e. x.

More generally, as part of proving a.e. convergence theorems on some $L_p(X, \beta, m)$, it is desirable to have weaker conditions on μ that guarantee at least that $T_{\mu^n} f$ converges a.e. for all $f \in S$ where S is dense in $L_p(X, \beta, \mu)$. Here are two results like this. See also Lin [20] for a proof and use of the next lemma. The assumption that G is abelian is not used in the next lemma.

2.9. LEMMA. *If μ is adapted and $h \in L_2(X, \beta, m)$ with $T_\mu h = h$, then h is G-invariant.*

Proof. Here $h = T_\mu h = \int T_g h \, d\mu(g)$ where $\|T_g h\|_2 = \|h\|_2$ for all $g \in G$. Assume, without loss of generality, $\|h\|_2 = 1$. Then $1 = \|h\|_2^2 = \langle T_\mu h, h\rangle = \int\langle T_g h, h\rangle \, d\mu(g)$. But $|\langle T_g h, h\rangle| \leq 1$. So for μ a.e. g, $\langle T_g h, h\rangle = 1$. That is, $T_g h = h$ for μ a.e. $g \in G$. Since μ is adapted, h is G-invariant. $\square$

The following also appears in Blum and Reich [7] and Tempelman [34], [35], 6.2.1, in greater generality.

2.10. PROPOSITION. *Suppose G is an abelian group and μ is strictly aperiodic. Then $\lim_{n\to\infty} T_\mu^n f(x)$ exist a.e. for all f in a subspace S which is dense in $L_2(X, \beta, m)$.*

Proof. Again, it suffices to consider a direct integral decomposition $T_\mu = \int \oplus \hat{\mu}(\gamma)\,\mathrm{Id}\, dF(\gamma)$ where (Λ, D) is $(\hat{G}, \beta_{\hat{G}})$. Consider a function $f = \int \oplus f_\gamma\, dF(\gamma) \in \int \oplus H_\gamma\, dF(\gamma)$ where $f_\gamma = 0$ for $\gamma \notin K$ and K is a compact subset of $\hat{G}$ with $1 \notin K$. By continuity of $\hat{\mu}$,

$$\sup\{|\hat{\mu}(\gamma)|: \gamma \in K\} = 1 - \varepsilon \quad \text{for some } \varepsilon > 0.$$

So

$$\begin{aligned} \|T_\mu^n f\|_2^2 &= \int |\hat{\mu}(\gamma)^n|^2 \|f_\gamma\|^2\, dF(\gamma) \\ &\le (1-\varepsilon)^{2n} \int \|f_\gamma\|^2\, dF(\gamma) \\ &= (1-\varepsilon)^{2n} \|f\|_2^2. \end{aligned}$$

Hence, $\sum_{n=1}^\infty \|T_\mu^n f\|_2^2$ converges and so $\sum_{n=1}^\infty |T_\mu^n f|^2$ is in $L_1(X, \beta, m)$. Thus, $\lim_{n\to\infty} T_\mu^n f(x) = 0$ a.e.. But the space of such functions f is dense in

$$\{f_0 \in L_2(X, \beta, \mu): \langle f_0, h\rangle = 0 \text{ for all } G\text{-invariant functions } h\}.$$

Indeed, if $\int \langle h_\gamma, f_\gamma \rangle\, dF(\gamma) = 0$ for all such f, then $h_\gamma = 0$ for F a.e. γ, except for $\gamma = 1$. So

$$T_\mu h = \int \oplus \hat{\mu}(\gamma) h_\gamma\, dF(\gamma) = \int \oplus h_\gamma\, dF(\gamma) = h.$$

But μ is strictly aperiodic and hence adapted. So h is G-invariant by Lemma 2.9.

This gives a dense subspace S of $L_2(X, \beta, m)$, namely $\{f_1 + f_2$: f_1 is G-invariant, $(f_2)_\gamma = 0$ for $\gamma \notin K$, for some $K \subset \hat{G}$, K compact, $1 \notin K\}$, on which $T_\mu^n f$ converges a.e. □

2.11. *Remark.* The conclusion in 2.8 is stronger than the one in 2.10 in that in 2.8 we get a dense subspace of any $L_p(X, \beta, m)$, $1 \le p < \infty$.

This is the analogous theorem for operators.

2.12. PROPOSITION. *Let T be a normal contraction operator on $L_2(X, \beta, m)$ and suppose that the resolution of the identity E corresponding to T has $E\{z: |z| = 1, z \ne 1\} = 0$. Then there is a dense subspace of $L_2(X, \beta, m)$ such that $\lim_{n\to\infty} T^n f(x)$ exists a.e. for all $f \in S$.*

2.13. *Remark.* (a) Sometimes it is clear that $T = T_\mu$ satisfies the hypotheses of 2.12. For example, if $G = Z$ and μ is strictly aperiodic, then the spectral mapping theorem shows $sp(T_\mu) \cap \{z\colon |z| = 1\} = \{1\}$, so trivially 2.12 holds because $E\{z\colon |z| = 1, z \neq 1\} = E(\emptyset) = 0$. However, if $G = R$, $\mu = \frac{1}{3}(\delta_a + \delta_b + \delta_c)$, where $a, b, c, 1$ are rationally independent, then μ is strictly aperiodic but $sp(T_\mu) = \{z\colon |z| \leq 1\}$.

(b) By 1.9 it is easy to see that $sp(T_\mu) \cap \{z\colon |z| = 1\} = \{1\}$ if and only if for any open neighborhood W of 1, $\sup_{\gamma \notin W} |\hat{\mu}(\gamma)| < 1$. It would be interesting to characterize when $E_T\{z\colon |z| = 1, z \neq 1\} = 0$ in a similar fashion. But in any case, we have Theorem 2.10 for strictly aperiodic measures.

The other part of getting a.e. convergence theorems for $T_\mu^n f$, $f \in L_p(X, \beta, m)$, is the existence of a maximal inequality. There are occasions when such a result is true along a suitable subsequence of $\{1, 2, 3, \ldots\}$. This was observed for discrete abelian groups in [3], [4], [5]. More generally, we have this result.

2.14. PROPOSITION. *Suppose G is an abelian group and μ is strictly aperiodic. Assume that for some compact neighborhood K of e, $\sup_{\gamma \notin K} |\hat{\mu}(\gamma)| < 1$. Then there is a subsequence $(n_s\colon s = 1, 2, 3, \ldots)$ such that $\lim_{s \to \infty} T_\mu^{n_s} f(x)$ exists a.e. for all $f \in L_2(X, \beta, m)$.*

Proof. Proposition 1.9 and the above guarantee that

$$sp(T_\mu) \cap \{z\colon |z| = 1\} = \{1\}.$$

Since T_μ is a normal contraction, this result is in [4]. □

If μ is adapted and $\lim_{n \to \infty} \|\mu^{n+1} - \mu^n\|_1 = 0$, then μ is strictly aperiodic because

$$\sup_{\gamma \in \hat{G}} |\hat{\mu}(\gamma)|^n |(\hat{\mu}(\gamma) - 1)| \leq \|\mu^{n+1} - \mu^n\|_1 \to 0$$

as $n \to \infty$. So if $\gamma \neq 1$, $\hat{\mu}(\gamma) \neq 1$ and $\hat{\mu}(\gamma)^n \to 0$ as $n \to \infty$ i.e. $|\hat{\mu}(\gamma)| < 1$ for $\gamma \neq 1$. But moreover, whenever $\sup_{\gamma \in \hat{G}} |\hat{\mu}(\gamma)|^n |\hat{\mu}(\gamma) - 1| \to 0$ as $n \to \infty$, we have $|\hat{\mu}(\gamma)|$ bounded away from 1 if $\hat{\mu}(\gamma)$ is bounded away from 1 i.e. $\mathrm{cl}_C\{\hat{\mu}(\gamma)\colon \gamma \in \hat{G}\} \cap \{z\colon |z| = 1\} = \{1\}$. Hence, these special types of strictly aperiodic measures also satisfy subsequence theorems by the same proof as the one for 2.14.

2.15. PROPOSITION. *If G is abelian and a probability measure $\mu \in M(G)$ has*

$$\sup_{\gamma \in \hat{G}} |\hat{\mu}(\gamma)|^n |\hat{\mu}(\gamma) - 1| \to 0$$

as $n \to \infty$, then there is an increasing subsequence $(n_s\colon s = 1, 2, 3, \ldots)$ such that $T_\mu^{n_s} f(x)$ converges a.e. for all $f \in L_2(X, \beta, m)$. □

2.16. *Remark.* (a) If $\mu \ll \lambda_G$ and strictly aperiodic, then the conditions of 2.14 hold. However, μ can fail to satisfy 2.14 and still satisfy 2.15. For example, let $\mu = \frac{1}{3}(\delta_0 + \delta_a + \delta_b)$. If a, b are rationally independent, then μ is adapted and strictly aperiodic. Also, $\|\mu^{n+1} - \mu^n\|_1 \to 0$ since μ^2 and μ are not mutually singular, by Foguel's theorem [14]. This norm condition is stronger than what is required for 2.15. However, by Kronecker's lemma, for a sequence (γ_s), $\lim_{s\to\infty}\gamma_s = \infty$, $\lim_{s\to\infty}\hat{\mu}(\gamma_s) = 1$. So the conditions of 2.14 fail.

(b) The example μ in Rudin [31] is a symmetric probability measure with

$$\|\Sigma_{s=1}^N \alpha_s \mu^s\|_1 = \Sigma_{s=1}^N |\alpha_s|$$

for all $\alpha_1, \ldots, \alpha_N \in \mathbf{C}$. Let $\nu = \mu^2$. Then $\|\nu^{n+1} - \nu^n\|_1 = 2$ for all $n \geq 1$, but

$$\sup_{\gamma\in\hat{G}} |\hat{\nu}(\gamma)|^n |1 - \hat{\nu}(\gamma)| \leq \sup_{0\leq r\leq 1} r^n(1-r) \leq 1/n$$

because $\hat{\nu}(\gamma) \in [0, 1]$ for all $\gamma \in \hat{G}$. Thus, the hypotheses of 2.15 hold, but $\|\nu^{n+1} - \nu^n\|_1$ does not go to zero. This is another reason to state 2.15 in this form. It would also be good to remark that the operator T_ν does have $\|T_\nu^{n+1} - T_\nu^n\|_2 \leq C/n$; indeed, this is true for any T which is normal and has spectrum contained in a proper Stolz region. See [4].

One can also obtain some subsequence theorems for all L_p, $1 < p < \infty$, by using the method in [3]. Namely, if G is $R^k \oplus Z^l$ for some $k, l \geq 0$, and μ is as in 2.14 or 2.15 then there exists a subsequence (n_s) such that $\lim_{s\to\infty} T_\mu^{n_s} f(x)$ exists a.e. for all $f \in L_p(X, \beta, m)$, $1 < p < \infty$. Since we do not see how to prove this more abstractly in abelian groups in general, the proof (which is identical to the one in [3]) will be left out here.

2.17. *Remark.* If G is a discrete abelian group, then strict aperiodicity is the only assumption needed in 2.14. However, in general strict aperiodicity is not sufficient. Moreover, without additional information about the behavior of $\hat{\mu}(\gamma)$ for γ near 1, there may be no way to replace (n_s) by $(1, 2, 3, \ldots)$. See [5].

With regards to going beyond norm theorems to get pointwise results, when only strict aperiodicity is assumed, the following is an example of the difficulties that can arise.

2.18. THEOREM. *Let a, b, c be real numbers such that 1, a, b and c are rationally independent. Let $\{T_t\colon t \in R\}$ be an aperiodic ergodic flow on a* *probability space* (X, β, m). *Let* $\mu = \frac{1}{3}(\delta_a + \delta_b + \delta_c)$. *Then for any increasing sequence* (n_s), *there is* $f \in L_\infty(X, \beta, m)$ *such that* $\lim_{s\to\infty} T_\mu^{n_s} f(x)$ *does not exist a.e.*.

Before proving this result, let us make some observations. First, μ is strictly aperiodic, but $\|\mu^{n+1} - \mu^n\|_1 = 2$ for all $n \geq 1$. Also, even $\|T_\mu^{n+1} - T_\mu^n\|_2 = 2$ for all $n \geq 1$ here. Second, $\{\hat{\mu}(\gamma)\colon \gamma \in R\}$ is dense in $\{z\colon |z| \leq 1\}$ by Kronecker's lemma. So the conditions of 2.14 and 2.15 both fail. However, because of strict aperiodicity, there is a norm convergence result by 2.4.

Finally, because of the norm identification in 2.4, 2.18 gives a μ such that the averages

$$\frac{1}{N}\sum_{k=1}^{N} T_\mu^k f$$

converge a.e. and in L_p-norm to $P_I f$ for all $f \in L_p(X, \beta, m)$, $1 \leq p < \infty$. Nonetheless, the chaotic behavior of $sp(T_\mu)$ causes there to be no (subsequence) theorem for a.e. convergence of iterates of T_μ in this case.

Proof 2.18. Because the flow is ergodic, the averages

$$A_N f = \frac{1}{N}\sum_{k=1}^{N} T_\mu^k f$$

converge in L_2-norm to $\int f\,dm$ for all $f \in L_2(X, \beta, m)$. Hence, to prove the result only requires denying the finiteness of the L_2-metric entropy for $(T_\mu^{n_s}\colon n \geq 1)$, by Bourgain [8]. See also Rosenblatt [29].

Choose $M \geq 1$ and (σ_{ij}), $i = 1, \ldots, 2^M$, $j = 1, \ldots, M$, with each $\sigma_{ij} = \pm 1$, such that column $(\sigma_{ij}\colon i \geq 1)$ consists of half $+1$ and half -1, and the columns $(\sigma_{ij}\colon i \geq 1)$ are independent of each other for distinct j. In Rosenblatt [29] it was shown that for all $\alpha > 0$, there is $r(\alpha) \geq 0$ such that if (η_k) is lacunary with $\eta_{k+1}/\eta_k \geq r(\alpha)$, and $s_k = \pm 1$, then there exists γ, $|\gamma| = 1$, such that $|\gamma^{\eta_k} - s_k| \leq \alpha$ for $k \geq 1$. Choose a subsequence $(n_{m_k}\colon k \geq 1)$ with $n_{m_{k+1}}/n_{m_k} \geq r(\alpha)$ for all $k \geq 1$. Choose γ_i, $|\gamma_i| = 1$, $|\gamma_i^{n_{m_j}} - \sigma_{ij}| \leq \alpha$ for $i = 1, \ldots, 2^M$, $j = 1, \ldots, M$. Now choose real numbers $r_i, r_{i+1} \geq r_i + 1$, $i = 1, \ldots, M$, with $|\hat{\mu}(r_i) - \gamma_i| \leq \varepsilon_0$. This is possible by the rational independence of $1, a, b, c$. By suitable choices of α and ε_0, we see that we can arrange $|\hat{\mu}^{n_{m_j}}(r_i) - \sigma_{ij}| \leq \varepsilon$ for $i = 1, \ldots, 2^M$, $j = 1, \ldots, M$.

Using Lind [21], the aperiodicity of $(T_t\colon t \in R)$ shows that for all $\varepsilon > 0$, there is a measure-theoretic copy of $(F \times [L, L], \mu_F \times \lambda_R)$ in (X, m) where μ_F is a positive measure on F with $m(X \setminus (F \times [-L, L])) < \varepsilon$, so that $T_t(x, r) = (x, r + t)$ for $(x, r) \in F \times [-L, L]$ with $t + r \in [-L, L]$ too. It follows that

$$\frac{1}{2L} \geq \mu_F(F) \geq \frac{1-\varepsilon}{2L}.$$

Define $\chi_r(x,t) = \exp(-2\pi i r t)$ for $(x,t) \in F \times [-L, L]$, and $\chi_r(x,t) = 0$ otherwise. Here $\chi_r(t)$ will also denote $\exp(-2\pi i r t)$, for any $t \in R$. Let

$$f(x,t) = \frac{1}{\sqrt{2^M}} \sum_{l=1}^{2^M} \chi_{r_l}(x,t).$$

This f has L_2-norm which can be estimated by

$$\|f\|_2^2 = m(F \times [-L, L]) + \frac{1}{2^M} \sum_{\substack{l,l'=1 \\ l \neq l'}}^{2^M} \int \chi_{r_l} \overline{\chi_{r_{l'}}}\, dm$$

and

$$\begin{aligned} \left| \int \chi_{r_l} \overline{\chi_{r_{l'}}}\, dm \right| &\leq \frac{1}{2L} \left| \int_{-L}^{L} e^{-2\pi i r_l t} e^{+2\pi i r_{l'} t}\, dt \right| \\ &\leq \frac{2}{2L \cdot 2\pi |r_l - r_{l'}|} \leq \frac{1}{2\pi L} \end{aligned}$$

since $|r_l - r_{l'}| \geq 1$ for $l \neq l'$. Hence

$$\begin{aligned} \|f\|_2^2 &\leq 1 + \frac{1}{2^M} 2^M (2^M - 1) \frac{1}{2\pi L}. \\ &= 1 + \frac{2^M - 1}{2\pi L} \end{aligned}$$

Also, for $(x,t) \in F \times [-L, L]$, such that $s + t \in [-L, L]$ for all $s \in \operatorname{supp}(\mu^n)$,

$$\begin{aligned} T_\mu^n f(x,t) &= \frac{1}{\sqrt{2^M}} \sum_{l=1}^{2^M} \sum_s \mu^n(s) \chi_{r_l}(x, s+t) \\ &= \frac{1}{\sqrt{2^M}} \sum_{l=1}^{2^M} \sum_s \mu^n(s) \exp(-2\pi i r_l (s+t)) \\ &= \frac{1}{\sqrt{2^M}} \sum_{l=1}^{2^M} \hat{\mu}^n(r_l) \chi_{r_l}(t). \end{aligned}$$

Hence, there is a constant K, depending only on M, such that if $|t| \leq L - K$, $x \in F$, and $k = 1, \ldots, M$, then

$$T_\mu^{n_{m_k}} f(x,t) = \frac{1}{\sqrt{2^M}} \sum_{l=1}^{2^M} \hat{\mu}^{n_{m_k}}(r_l) \chi_{r_l}(t).$$

But then for $k \neq k'$,

$$
\begin{aligned}
&\|T_\mu^{n_{m_k}} f - T_\mu^{n_{m_{k'}}} f\|_2^2 \\
&\quad \geq \int_F \int_{-L+K}^{L-K} \left|(T_\mu^{n_{m_k}} f - T^{n_{m_{k'}}} f)(x,t)\right|^2 dt\, d\mu_F(x) \\
&\quad = \int_F \int_{-L+K}^{L-K} \frac{1}{2^M} \left| \sum_{l=1}^{2^M} (\hat{\mu}^{n_{m_k}}(r_l) - \hat{\mu}^{n_{m_{k'}}}(r_l)) \chi_{r_l}(t) \right|^2 dt\, d\mu_F(x) \\
&\quad \geq \frac{1-\varepsilon}{2L} \int_{-L+K}^{L-K} \frac{1}{2^M} \left| \sum_{l=1}^{2^M} (\hat{\mu}^{n_{m_k}}(r_l) - \hat{\mu}^{n_{m_{k'}}}(r_l)) \chi r_l(t) \right|^2 dt.
\end{aligned}
$$

But for $l \neq l'$,

$$
\left| \int_{-L+K}^{L-K} \chi_{r_l}(t) \overline{\chi_{r_{l'}}(t)}\, dt \right| \leq \frac{2}{2\pi |r_l - r_{l'}|} \leq \frac{1}{\pi}
$$

since $|r_l - r_{l'}| \geq 1$ for $l \neq l'$. So expanding the integrand above gives, for $k \neq k'$,

$$
\begin{aligned}
&\|T_\mu^{n_{m_k}} f - T_\mu^{n_{m_{k'}}} f\|_2^2 \\
&\quad \geq \frac{(1-\varepsilon)(2L-2K)}{2L \cdot 2^M} \sum_{l=1}^{2^M} |\hat{\mu}^{n_{m_k}}(r_l) - \hat{\mu}^{n_{m_{k'}}}(r_l)|^2 \\
&\qquad - \frac{1-\varepsilon}{2L \cdot 2^M} 2^M (2^M - 1) \frac{4}{\pi}.
\end{aligned}
$$

If L is large enough, and $\varepsilon > 0$ small enough, for $k \neq k'$,

$$
\begin{aligned}
&\|T_\mu^{n_{m_k}} f - T_\mu^{n_{m_{k'}}} f\|_2^2 \\
&\quad \geq \frac{(1-\varepsilon)(2L-2K)}{2L \cdot 2^M} \sum_{l=1}^{2^M} |\sigma_{lk} - \sigma_{lk'}|^2 \\
&\qquad - \frac{(1-\varepsilon)(2L-2K)}{2L \cdot 2^M} \sum_{l=1}^{2^M} 12\varepsilon - \frac{(1-\varepsilon)(2^M - 1)}{2\pi L} \\
&\quad \geq \frac{1}{2^{M+1}} \sum_{l=1}^{2^M} |\sigma_{lk} - \sigma_{lk'}|^2 - \frac{1}{5} \\
&\quad \geq \frac{1}{2^{M+1}} (4 \cdot 2^{M-1}) - \frac{1}{5} \geq \frac{4}{5}.
\end{aligned}
$$

At the same time, if L is large enough and $\varepsilon > 0$ is small enough, we can have $\frac{3}{2} \geq \|f\|_2^2$.

The conclusion is that for any $M \geq 1$ there exists $f \in L_2(X, \beta, m)$, $\|f\|_2 = 1$ and some $n_{m_1}, \ldots, n_{m_M}$ such that $\|T_\mu^{n_{m_k}} f - T_\mu^{n_{m_{k'}}} f\|_2 \geq \sqrt{8/15}$ for all $k, k' = 1, \ldots, M$, $k \neq k'$. This shows that the metric entropy of T_μ^n is infinite. □

2.19. *Remark.* (a) The same conclusion as 2.18 is true if we just know that some non-trivial arc in $\{z\colon |z| = 1\}$ is a subset of $\mathrm{cl}_{\mathbf{C}}\{\hat{\mu}(r)\colon r \in R\}$.

(b) The example in 2.18 is important because all previously studied cases of the use of convolution powers to get ergodic theorems had the property that if there were an L_p-norm theorem, then there was at least an a.e. convergence theorem along a subsequence of powers (n_s). See Bellow, Jones, Rosenblatt [3].

(c) There should be an analogous theorem for a normal operator T with $sp(T)$ containing an arc in the unit circle.

(d) An inspection of the proof of 2.18 shows that the constant δ in the entropy calculation can be, instead of $\sqrt{8/15}$, as chose to $\sqrt{2}$ as we like. For geometrical reasons, this is certainly optimal. It follows, analogously to Rosenblatt [29], that the powers T_μ^r are δ_0-sweeping out for some δ_0. Actually, there is a universal constant C such that $\delta_0 \geq C\delta$ where δ is as above, so that δ_0 does not depend on the dynamical system.

One final remark about convergence of powers for abelian groups in the positive direction.

2.20. THEOREM. *Let G be an abelian group and suppose μ is strictly aperiodic with $\{\hat{\mu}(\gamma)\colon \gamma \in \hat{G}\}$ contained in a proper Stolz region. Then for all $f \in L_p(X, \beta, m)$, $1 < p < \infty$, $T_\mu^n f$ converges a.e. and in L_p-norm to $P_I f$.*

Proof. The normal operator T_μ has $sp(T_\mu) \subset \mathrm{cl}_{\mathbf{C}}\{\hat{\mu}(\gamma)\colon \gamma \in \hat{G}\}$. So Bellow, Jones, Rosenblatt [4] gives this general result. □

2.21. *Remark.* The condition in 2.20 can be phrased as

$$\sup_{\substack{\gamma \in \hat{G} \\ \gamma \neq 1}} \frac{|1 - \hat{\mu}(\gamma)|}{1 - |\hat{\mu}(\gamma)|} < \infty.$$

This generalizes to abelian groups the analogous result in [4], [5]. In particular, if $G = Z^m$, and μ is a probability measure on G with finite second moment, $\sum_{\mathbf{k} \in Z^m} \|\mathbf{k}\|^2 \mu(\mathbf{k}) < \infty$, then the expectations

$$\sum \{k_i \mu(\mathbf{k})\colon \mathbf{k} \in Z^m, \mathbf{k} = (k_1, \ldots, k_m), k_i \in Z\}$$

all being zero is necessary and sufficient for

$$\sup_{\substack{\gamma \in T^m \\ \gamma \neq 1}} \frac{|1 - \hat{\mu}(\gamma)|}{1 - |\hat{\mu}(\gamma)|} < \infty.$$

This generalizes the result in [5] for $G = Z$.

3. General locally compact groups

In Section 2, we saw the principle that if μ is a strictly aperiodic measure on an abelian group G, and (X, β, m) is a probability space, then T_μ^n satisfies a norm convergence theorem for all $f \in L_p(X, \beta, m)$, $1 \leq p < \infty$; but $T_\mu^n f$ may or may not converge a.e. for $f \in L_p(X, \beta, m)$ even for a subsequence of powers, depending on more delicate features of the behavior of $\hat{\mu}$. In general locally compact groups, some of these positive results still hold, but the difficulties with getting simple general theorems become even greater.

Perhaps the first principle to observe is the following very general one. This is stated in Tempelman [33], see also Tempelman [34], [35].

3.1. THEOREM. *If μ is an adapted probability measure on G, and (X, β, m) is a probability space then for all $f \in L_p(X, \beta, m)$, $1 \leq p < \infty$, the averages*

$$\frac{1}{n} \sum_{n=1}^{N} T_\mu^n f(x)$$

converge a.e. and in L_p-norm to the invariant projection $P_I f$ of f on the G-invariant functions in $L_p(X, \beta, m)$.

Proof. By Lemma 2.6, which did not require commutativity of G, $\{T_\mu f - f: f \in L_2(X, \beta, m)\}$ is a dense subspace of $I^\perp$, where I is the subspace of G-invariant functions in $L_2(X, \beta, m)$. Hence,

$$\lim_{N \to \infty} \left\| \frac{1}{N} \sum_{n=1}^{N} T_\mu^n f - P_I f \right\|_2 = 0$$

for all $f \in L_p(X, \beta, m)$, $1 \leq p < \infty$. But also by the Dunford-Schwartz theorem, $1/N)\Sigma_{n=1}^{N} T_\mu^n f(x)$ converges a.e. for all $f \in L_1(X, \beta, m)$. □

3.2. *Remark.* This result answers a question in Milnes and Paterson [23] in that it shows how there always exists, what they call, ergodic sequences in a group.

To obtain norm and a.e. convergence theorems for T_μ^n by itself, further conditions on μ are required. One example of a general result of this type is that for any σ-compact amenable group there exists $\varphi \in L_1(G, \beta_G, \lambda_G)$, such that $d\mu = \varphi\, d\lambda_G$ is a probability measure with $\lim_{n\to\infty} \|\mu^n * h\|_1 = 0$ for all $h \in L_1(G, \beta_G, \lambda_G)$, $\int h\, d\lambda_G = 0$. See Rosenblatt [27] and Kaimanovich and Vershik [19]. These give examples of when the following occurs. Also compare this result with Foguel [15] which shows that in the abelian case, the hypotheses of 2.4 and 3.3 are the same.

3.3. THEOREM. *Assume G is a σ-compact metric group. Let μ be a probability measure on the group such that $\lim_{n\to\infty} \|\mu^n * h\|_1 = 0$ for all $h \in L_1(G, \beta_G, \lambda_G)$, $\int h\, d\lambda_G = 0$. Let (X, β, m) be a probability space. Assume the representation $g \mapsto T_g$ on $L_2(X, \beta, m)$ is continuous. Then for any $f \in L_p(X, \beta, m)$, $1 \le p < \infty$, $\lim_{n\to\infty} T_\mu^n f = P_I f$ in L_p-norm.*

Proof. It suffices to show that $\lim_{n\to\infty} \|T_\mu^n f\|_2 = 0$ for all $f \in L_2(X, \beta, m)$ of the form $T_h f_0$ for some $h \in L_1(G, \beta, \lambda_G)$, $\int h\, d\lambda_G = 0$, and some $f_0 \in L_2(X, \beta, m)$. Indeed, if $f_1 \in L_2(X, \beta, m)$, and $\langle f_1, T_h f_0\rangle = 0$ for all such h and f_0, then $(T_{h^*}) f_1 = 0$ for all such h where h^* is defined by $h^*\, d\lambda_G = (h\, d\lambda_G)^*$. We claim this means f_1 is G-invariant. Choose $(h_n) \in L_1(G, \beta_G, \lambda_G)$, $\int_G h_n\, d\lambda_G = 1$, h_n of compact support and $\operatorname{supp} h_n \supset \operatorname{supp} h_{n+1}$, $\bigcap_{n=1}^\infty \operatorname{supp} h_n = \{e\}$. Then

$$\int h_n(g) h(g)\, d\lambda_G(g) \to h(e)$$

for all h continuous on G. But for any $f_2 \in L_2(X, \beta, m)$, $g \mapsto \int T_g f(x) \overline{f_2(x)}\, dm(x)$ is continuous. Also,

$$T_{h_n} T_{g_0} - T_{h_n} = T_{h_n * \delta_{g_0}} - T_{h_n} = T_{h_n * \delta_{g_0} - h_n}$$

and

$$\int h_n * \delta_{g_0}(z)\, d\lambda_G(z) = \int h_n\left(z g_0^{-1}\right) d\lambda_G(z) = \int h_n(z)\, d\lambda_G(z) = 1.$$

Hence, $h_n * \delta_{g_0} - h_n$, is mean zero and

$$\begin{aligned} 0 &= \left\langle \left(T_{h_n * \delta_{g_0} - h_n}\right) f_1, f_2 \right\rangle \\ 0 &= \left\langle \left(T_{h_n} T_{g_0} - T_{h_n}\right) f_1, f_2 \right\rangle \\ &= \left\langle T_{h_n}\left(T_{g_0} f_1\right), f_2 \right\rangle - \left\langle T_{h_n} f_1, f_2 \right\rangle. \end{aligned}$$

But

$$\begin{aligned}&\left\langle T_{h_n}(T_{g_0}f_1), f_2\right\rangle\\ &\quad = \int_G h_n(g)\int_X T_g(T_{g_0}f_1)(x)\overline{f_2(x)}\,dm(x)\,d\lambda_G(g)\\ &\quad \to \int_X (T_{g_0}f_1)\overline{f_2}\,dm \quad \text{as } n\to\infty.\end{aligned}$$

So letting $n\to\infty$, for all $f_2 \in L_2(X,\beta,m)$,

$$0 = \left\langle T_{g_0}f_1, f_2\right\rangle - \left\langle f_1, f_2\right\rangle.$$

That is, $T_{g_0}f_1 = f_1$ for all $g_0 \in G$.

The conclusion is that $\{T_h f_0\colon\ h \in L_1(G,\beta_G,\lambda_G),\ \int h\,d\lambda_G = 0,\ f_0 \in L_2(X,\beta,m)\}$ spans a subspace $S \subset L_2(X,\beta,m)$ with $S^\perp \subset I$. That is, $S \supset I^\perp$. Thus, if $\lim_{n\to\infty}\|T_\mu^n f\|_2 = 0$ for all $f \in S$, then the theorem is proved. But if $f = T_h f_0$, then $T_\mu^n f = T_{\mu^n * h} f_0$. Now $\lim_{n\to\infty}\|\mu^n * h\|_1 = 0$. So $\lim_{n\to\infty}\|T_{\mu^n * h} f_0\|_2 = 0$ too. □

This gives a fairly general norm convergence theorem. But with conditions on μ relating to the behavior of T_μ^γ, γ a continuous irreducible representation of G, there are similar results. See also Tempelman [34], [35] for a general spectral criterion for norm convergence which gives this proposition.

3.4. Proposition. *Assume G is a σ-compact metric group. Let $\mu \in M(G)$ be a probability measure. Suppose for all $\lambda \neq \mathrm{Id}$, $\|T_\mu^\lambda\| < 1$. Let (X,β,m) be a probability space. Then for all $f \in L_p(X,\beta,m)$, $1 \le p < \infty$, $\lim_{n\to\infty}\|T_\mu^n f - P_I f\|_p = 0$.*

Proof. We write the representation T on $I^\perp \subset L_2(X,\beta,m)$ as a direct integral which decomposes $T_\mu = \int_\Lambda \oplus\, T_\mu^\lambda\, dF(\lambda)$ where each λ is irreducible. Now consider $f = \int \oplus f^\lambda\, dF(\lambda)$ where $f^\lambda = 0$ if $\|T_\mu^\lambda\| \ge 1-\varepsilon$. Then

$$\begin{aligned}\|Tf\|_2^2 &= \int_\Lambda \left\|T_\mu^\lambda f^\lambda\right\|_2^2 dF(\lambda)\\ &\le (1-\varepsilon)^2\|f\|_2^2.\end{aligned}$$

So $\lim_{n\to\infty}\|T_\mu^n f\|_2 = 0$ for all such f. However, by the criterion in the theorem, such f span a dense subspace of $I^\perp$. □

3.5. *Remark.* It would be enough to assume $\|T_\mu^\lambda\| < 1$ for F a.e. λ. But since F is not specified, we have imposed the stronger hypothesis. The result in Derriennic and Lin [9] gives the conclusion above with different hypotheses.

The criterion of 3.4 turns out to be sometimes impossible to achieve. The best type of result would be one where a property of supp(μ) implies this condition. However, this is the best that can be said in this regard. See Fell [12], [13] and Dixmier [10] for background.

3.6. THEOREM. *Suppose μ is a strictly aperiodic probability measure on a σ-compact metric group G and λ is a continuous irreducible representation λ of G. Assume T on $L_2(X, \beta, m)$ is continuous. If $\|T_\mu^\lambda\| = 1$, then there exists a dense subgroup H and G such that the identity representation of H is weakly contained in λ in the discrete topology.*

Proof. If $\|T_\mu^\lambda\| = 1$ for some λ, then there exists $f_n \in H_\lambda$, $\|f_n\| = 1$, $\lim_{n\to\infty}\|T_\mu^\lambda f_n\| = 1$. But then

$$\|T_\mu^\lambda f_n\|_2^2 = \iint \langle T_g^\lambda f_n, T_h^\lambda f_n\rangle \, d\mu(g)\, d\mu(h)$$

$$= \iint \langle T_{h^{-1}g}^\lambda f_n, f_n\rangle \, d\mu(g)\, d\mu(h).$$

Now $|\langle T_{h^{-1}g}^\lambda f_n, f_n\rangle| \le 1$ for all h, $g \in G$. Hence, it must be that for some subsequence (f_{n_m}) and for $\mu \times \mu$ a.e. (g, h), $\lim_{m\to\infty}\langle T_{h^{-1}g}^\lambda f_{n_m}, f_{n_m}\rangle = 1$. The strict aperiodicity of μ shows that there is a dense subgroup H of G such that $\lim_{m\to\infty}\langle T_x^\lambda f_{n_m}, f_{n_m}\rangle = 1$ for all $x \in H$. That is, $\mathrm{Id}|_H$ is weakly contained in λ in $\hat{G}$ with the discrete topology. □

3.7. *Remark.* The same point makes it clear that if Id is weakly contained in λ, then for all probability measures $\mu \in M(G)$, $\|T_\mu^\lambda\| = 1$. So no condition on μ can get strict inequality as in Proposition 3.4 in this situation. Indeed, this is the case if one only knows Id is weakly contained in λ with the discrete topology.

3.8. COROLLARY. *Let G be a σ-compact metric nilpotent group and let μ be a strictly aperiodic measure on G. Assume (X, β, m) is a probability space and T is continuous on $L_2(X, \beta, m)$. Then for all $f \in L_p(X, \beta, m)$, $1 \le p < \infty$, $\lim_{n\to\infty}\|T_\mu^n f - P_I f\|_p = 0$.*

Proof. It suffices to prove that $\|T_\mu^\lambda\| < 1$ for all non-trivial continuous irreducible λ on G. By Theorem 3.6, otherwise there exists a dense subgroup H of G such that $\mathrm{Id}|_H$ is weakly contained in λ with the discrete topology. Since H is dense in G, $\lambda|_H$ is also an irreducible representation, $\lambda|_H \ne \mathrm{Id}|_H$.

But H is nilpotent and so an irreducible representation γ of H does not weakly contain $\mathrm{Id}|_H$. □

3.9. *Remark.* If G is nilpotent and μ is a spread-out strictly aperiodic measure then μ satisfies the conditions needed in 3.3. The corollary above shows that we can get this same conclusion, via a direct integral argument, without the spread-out condition. See Derriennic and Lin [9] for a more general result in the case that μ is spread-out.

Because of the above results, some discussion of the separation property that is needed would be worthwhile. In part, the point is that $\hat{G}$ is not generally Hausdorff (or even T_0 or T_1). See Fell [12], [13] and Bekka and Kaniuth [2]. If G is abelian, then $\hat{G}$ is the same in the Fell topology and in the dual topology. Hence, $\hat{G}$ is always Hausdorff. Also, if G is compact, then $\hat{G}$ is always discrete. R. Howe pointed out that the criterion needed above is true for nilpotent groups, even though $\hat{G}$ can fail to be Hausdorff in the nilpotent case.

However, it is clear that a condition like $\lambda \neq \mathrm{Id}$ not weakly containing Id will fail in general. In Yoshizawa [38], it is shown how a non-abelian free group has some $\lambda \in \hat{G}$ which weakly contain all irreducible representations. It is also easy to see that solvable groups may fail to have the needed property. R. Howe has also suggested this finitely-generated example of such a solvable group.

3.10. *Example.* Let M be a 2×2 matrix, e.g.,

$$\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix},$$

which is invertible over Z and has no eigenvalues of unit modulus. Let η: $Z \to \mathrm{Aut}(Z^2)$ be the homomorphism defined by $\eta(1) = M$. Then let G be the semi-direct product $Z^2 \times_\eta Z$. That is, G can be thought of as triples $((z_1, z_2), g^j)$, $z_1, z_2, j \in Z$, with the multiplication,

$$\big((z_1, z_2), g^j\big) \cdot \big((u_1, u_2), g^k\big) = \big((z_1, z_2) + M^j(u_1, u_2), g^{j+k}\big)$$

for all $z_i, u_i, j, k \in Z$. Then G is finitely generated by $((1,0), g^0), ((0,1), g^0)$, and $((0,0), g))$. It is also solvable since Z^2 is normal in G as written with $G/Z^2 = Z$.

Consider the representation λ of G on $l_2(Z^2)$ given as the regular representation induced by

$$\lambda\big((z_1, z_2), g^j\big)(u_1, u_2) = (z_1, z_2) + M^j(u_1, u_2).$$

Since G is solvable, and hence amenable, there exists an invariant mean on $l_\infty(Z^2)$ for this action. Using the Følner condition (see Rosenblatt [26] for a discussion of this idea due to Følner), there is a sequence (F_n) of finite sets in Z^2 such that for all $z_i, j \in Z$,

$$\lim_{n\to\infty} \frac{\operatorname{card}\big(\lambda(z_1, z_2, g^i)F_n \,\Delta F_n\big)}{\operatorname{card}(F_n)} = 0.$$

Let

$$f_n = \frac{1}{\sqrt{\operatorname{card}(F_n)}} 1_{F_n}.$$

Then $f_n \in l_2(Z^2)$, $\|f_n\|_2 = 1$, and $\lim_{n\to\infty}\|\lambda(g)f_n - f_n\|_2 = 0$ for all $g \in G$. Hence, λ weakly contains Id.

However $\lambda \neq$ Id, and we claim λ is irreducible. Indeed, if $S \subset l_2(Z^2)$ is a non-zero subspace, which is invariant under $\lambda(G)$, then we claim S is dense in $l_2(Z^2)$. Suppose $f \in S$, $f \neq 0$, and $F \in S^\perp$. Then by using $\lambda(Z^2 \times \{g^0\})$ as an action, one sees $f * \tilde{F} = 0$ as an element in $l_2(Z^2)$ where $*$ is convolution over Z^2. But then $0 = \widehat{f * \tilde{F}} = \hat{f}\hat{\tilde{F}}$ on $\widehat{Z^2} = T^2$. Now the invariance under M for S also gives $\hat{f} \cdot \hat{\tilde{F}} \circ \tau_M^j = 0$ for all $j \in z$ where τ_M is the automorphism of T^2 given by $M^T = M$, i.e.,

$$\tau_M(\gamma_1, \gamma_2) = \left(\gamma_1\gamma_2^2, \gamma_1^2\gamma_2^3\right)$$

in the case

$$M = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}.$$

Since M has no eigenvalues of modulus 1, τ_M is ergodic. Since $\hat{f} \neq 0$ on some set of positive measure, this ergodicity shows $\hat{\tilde{F}} = 0$ a.e. on T^2, i.e., $\tilde{F} = 0$ and so $F + 0$.

The example above seems to be well known, but we discuss it in such detail here because of the interesting mixture of ergodic theory and harmonic analysis involved in its construction. E. Kaniuth pointed out that this example is also a special case of a property of countable amenable discrete groups G such that $\{e\}$ is the only finite conjugacy orbit. For such groups, there is an irreducible representation λ such that λ weakly contains every other irreducible representation because λ is weakly equivalent to the regular representation.

It is also important to remark that the condition of 3.4 does also hold for some solvable, non-nilpotent groups. For example, if G is the motion group of the plane, then the conclusion of 3.6 is never the case, so 3.4 holds. This is also true for the higher dimensional motion groups. it remains to classify even the connected Lie groups which cannot have 3.6 occur. Since this separation property is so useful, but different than the types of separation properties of $\hat{G}$ that have already been studied, this would be a worthwhile project.

Finally, we should observe that the norm convergence theorem above for a general strictly aperiodic measure μ may hold well past the scope of where the norm criteria above can be used. For example, let G be the affine group $R \times_\eta R$ with the multiplication given by $\eta(r)(s) = e^r s$ for all $r, s \in R$. Then G is solvable, non-nilpotent, and the criterion of 3.4 fails. However, it can be shown that for any μ which is strictly aperiodic just on the factor group R, one has the conclusion of Theorem 3.4. It seems likely even that by considering induced representations, there is a norm convergence theorem based only on criterion on the supp(μ) for all connected solvable groups. Since there is also such a result for many of the semisimple groups, this would give impetus to proving such a general norm convergence result for all Lie groups. The details of this remain to be investigated.

Because of the problems mentioned above with norm techniques, it would be useful to state here this general fact. It also shows that the ensuing discussion about pointwise behavior of $T_\mu^n f$ can sometimes be completely resolved. See also Oseledec [25] and Tempelman [33]. This type of theorem is discussed further in Tempelman [34], [35], 4.1.5 and 6.6.1.

3.11. THEOREM. *Let $\mu \in M(G)$ be a symmetric strictly aperiodic probability measure. Let (X, β, m) be a probability space. Then $T_\mu^n f$ converges in norm to $P_I f$ for all $f \in L_p(X, \beta, m)$, $1 \leq p < \infty$, and $T_\mu^n f(x)$ converges for a.e. x if $f \in L_p(X, \beta, m)$, $1 < p \leq \infty$.*

Proof. Because μ is symmetric, T_μ is normal and by Stein [32], there is a strong maximal inequality

$$\left\| \sup_{n \geq 1} |T_\mu^n f| \right\|_p \leq C_p \|f\|_p \quad \text{if } 1 < p < \infty.$$

Let $J = \{f \in L_2(X, \beta, \mu): T_\mu f = -f\}$. If $f \in J$, then $T_{\mu^2} f = f$. But $\mu^2 = \mu^* * \mu$, which is adapted because μ is strictly aperiodic. So f is G-invariant and hence $f = 0$. Thus, the G-invariant functions I is exactly $\{f \in L_2(X, \beta, m): T_\mu f = f\}$. Now since T_μ is symmetric, $sp(T_\mu) \subset [-1, 1]$. So $I^\perp$ contains as a dense subspace

$$S = \left\{E_{T_\mu}(-1 + \varepsilon, 1 - \varepsilon)f \colon \varepsilon > 0, f \in L_2(X, \beta, m)\right\}.$$

Also, $\|T_\mu^n E_{T_\mu}(-1+\varepsilon, 1-\varepsilon)f\|_2 \le (1-\varepsilon)^n \|f\|_2$. Hence, $T_\mu^n f(x) \to 0$ a.e. for any $f \in S$. Thus, the maximal inequality shows that for all p, $1 < p \le \infty$, $T_\mu^n f(x)$ converges a.e. to $P_I f$. But then it follows that $T_\mu^n f$ converges in L_p-norm to $P_I f$ for all $f \in L_p(X, \beta, m)$ and all p, $1 \le p < \infty$. □

The argument used above, as far as it relates a.e. convergence on a dense subspace S, also applied in the context of Theorem 3.4.

3.12. THEOREM. *Assume G is a σ-compact metric group. Suppose μ is a probability measure such that for all $\lambda \neq \mathrm{Id}$, $\|T_\mu^\lambda\| < 1$. Then there is a dense subspace S of $L_2(X, \beta, m)$ such that $\lim_{n\to\infty} T_\mu^n f(x)$ exists a.e. x for all $f \in S$.*

Proof. This is argued identically to the argument used in the abelian case via direct integrals. □

Sometimes the criterion of 3.12 can be improved enough to get a convergence result. Here is an example of such a theorem.

3.13. THEOREM. *Let G be a compact metric group and let $\mu \in M(G)$ be a probability measure such that $\sup_{\gamma \in \hat{G},\ \gamma \neq \mathrm{Id}} \|T_\mu^\gamma\| < 1$. Then for all $f \in L_p(G, \beta_G, \lambda_G)$, $1 < p \le \infty$, $T_\mu^n f(x)$ converges a.e. to $\int f\, d\lambda_G$. Also, $T_\mu^n f$ converges in L_p-norm to $\int f\, d\lambda_G$ for all $f \in L_p(G, \beta_G, \lambda_G)$, $1 \le p < \infty$.*

Proof. Only the a.e. convergence has not been proved. By a direct integral argument, if

$$L_2^0(G, \beta_G, \lambda_G) = \left\{ f \in L_2(G, \beta_G, \lambda_G) : \int f\, d\lambda_G = 0 \right\},$$

then

$$\|T_\mu\|_{L_2^0} \le \sup_{\substack{\gamma \in \hat{G} \\ \gamma \neq \mathrm{Id}}} \left\| T_\mu^\gamma \right\| < 1.$$

Since $\|T_\mu\|_{L_1^0} \le 1$, by interpolation (see [28]), $\|T_\mu\|_{L_p^0} \le 1 - \delta_p$, for some $\delta_p > 0$, if $1 < p < \infty$. This gives the strong convergence $\sum_{n=1}^\infty \|T_\mu^n f\|_q^p < \infty$ for all $f \in L_p^0(G, \beta_G, \lambda_G)$, $1 < p < \infty$. Hence, $\lim_{n\to\infty} T_\mu^n f(x) = 0$ a.e. for these f. This proves the theorem. □

3.14. *Examples.* For example, if G is abelian, then any $f \in L_1(G, \beta_G, \lambda_G)$ with support not contained in a proper closed coset of G will do. There are also examples though where μ is continuous and singular to λ_G, but $\hat{\mu}$ vanishes at ∞ to which the above applies.

If the group G has a stronger representation property, then not only will 3.5 be satisfied but there is a uniformity too.

3.15. THEOREM. *Let G be a discrete group with Kahzdan's property T. Suppose μ is strictly aperiodic on G. Then* $\sup_{\gamma\in\hat{G},\gamma\neq\mathrm{Id}}\|T_\mu^\gamma\| < 1$. *Let (X,β,m) be a probability space. Then for all $f\in L_p(X,\beta,m)$, $1<p<\infty$, $\lim_{n\to\infty}T_\mu^n f(x)$ exists a.e. and is $P_I f$. There is also norm convergence to the same limit on each of $L_p(X,\beta,m)$, $1\le p<\infty$.*

Proof. If there exists $(\gamma_n)\subset\hat{G}\setminus\{\mathrm{Id}\}$ such that $\|T_\mu^{\gamma_n}\|\to 1$, then on $H=\oplus_{n=1}^\infty H_{\gamma_n}$, $\|T_\mu\|=1$. Hence, by the strict aperiodicity of μ, Id is weakly contained in the diagonal representation T of G on H given by $T_g(f_n)=(T_g^{\gamma_n}f_n)$ for all $(f_n)\in H$. By property T, then H contains a G-invariant vector (f_n) with $\Sigma_{n=1}^\infty\|f_n\|_{H\gamma_n}^2=1$. But then each f_n is G-invariant. So some γ_n is trivial, a contradiction. Now $\|T_\mu\|_{I^\perp}<1$ and so the rest of the argument proceeds identically to the one in 3.13. □

3.16. *Remark*. We see that what is being said above in part is that if G has property T (i.e., Id is isolated in $\hat{G}$), then, by a well-known aspect of property T, if Id is weakly contained in another (irreducible) representation U, then Id is actually a subrepresentation of U. So if U is irreducible too, then $U=\mathrm{Id}$. That is, when $\{\mathrm{Id}\}$ is open in the Fell topology, then $\mathrm{Id}\notin\mathrm{cl}_{\hat{G}}\{\gamma\}$ for any $\gamma\in\hat{G}$, $\gamma\neq\mathrm{Id}$.

Another weaker condition on $\|T_\mu^\gamma\|$ which gives an a.e. convergence theorem is one restricting the behavior of $\|T_\mu^\gamma\|$ as γ converges to Id.

3.17. THEOREM. *Suppose the probability measure $\mu\in M(G)$ has $\|T_\mu^\gamma\|<1$ for all $\gamma\in\hat{G}\setminus\{\mathrm{Id}\}$, and*

$$\sup_{\substack{\gamma\in\hat{G}\\ \gamma\neq\mathrm{Id}}}\frac{\|\mathrm{Id}-T_\mu^\gamma\|}{1-\|T_\mu^\gamma\|}<\infty.$$

Then for all $f\in L_p(X,\beta,m)$, $1<p<\infty$, $\lim_{n\to\infty}T_\mu^n f(x)$ exists a.e.

Proof. We know there is a dense subspace with a.e. convergence by 3.12. The only question is then if there is a maximal inequality. Consider first $p=2$. As in [4], [5], we can bound,

$$(1)\quad \left\|\sup_{n\geq 1}T_\mu^{n+1}f\right\|_2\leq\left\|\sup_{n\geq 1}\frac{1}{n}\sum_{k=1}^n T_\mu^k f\right\|_2+\left\|\left(\sum_{k=1}^\infty k\left|\left(T_\mu^{k+1}-T_\mu^k\right)f\right|^2\right)^{1/2}\right\|_2.$$

But

$$\begin{aligned}\left\| \left(\sum_{k=1}^{\infty} k \left| (T_\mu^{k+1} - T_\mu^k) f \right|^2 \right)^{1/2} \right\|_2^2 &= \sum_{k=1}^{\infty} k \int \left| (T_\mu^{k+1} - T_\mu^k) f \right|^2 dm \\ &= \sum_{k=1}^{\infty} k \int_\Lambda \left\| \left((T_\mu^\lambda)^{k+1} - (T_\mu^\lambda)^k \right) f_\lambda \right\|^2 dF(\lambda) \\ &\le \int_\Lambda \sum_{k=1}^{\infty} k \| T_\mu^\lambda \|^{2k} \left\| I - T_\mu^\lambda \right\|^2 \|f_\lambda\|^2 \, dF(\lambda).\end{aligned}$$

But for $\lambda \ne \mathrm{Id}$,

$$\sum_{k=1}^{\infty} k \left\| T_\mu^\lambda \right\|^{2k} \left\| \mathrm{Id} - T_\mu^\lambda \right\|^2 \le \left(\frac{\left\| \mathrm{Id} - T_\mu^\lambda \right\|}{1 - \left\| T_\mu^\lambda \right\|} \right)^2 .$$

Hence, since the series is zero for $\lambda = \mathrm{Id}$,

$$\left\| \left(\sum_{k=1}^{\infty} k \left| (T_\mu^{k+1} - T_\mu^k) f \right|^2 \right)^{1/2} \right\|_2 \le \left(\sup_{\substack{\gamma \in \hat{G} \\ \gamma \ne \mathrm{Id}}} \frac{\left\| \mathrm{Id} - T_\mu^\gamma \right\|}{1 - \left\| T_\mu^\gamma \right\|} \right) \|f\|_2 . \tag{2}$$

Hence, (1) and (2) give a strong L_2 maximal inequality with the hypothesis of the theorem.

A similar calculation shows that the r-th difference operator $\Delta^r T_\mu^n$, satisfies

$$\left\| \sup_{n \ge r} n^r \left| \Delta^r T_\mu^n f \right| \right\|_2 \le C_r \|f\|_2$$

for some constant C_r. The rest of the argument, to get a strong L_p maximal inequality, $1 < p < \infty$, proceeds as in Stein [32] by using his complex interpolation theorem. Thus, there is always a strong maximal inequality in $L_p(X, \beta, m)$, $1 < p < \infty$. This is all that was needed to prove the theorem. □

3.18. *Remark.* It would be interesting, for μ adapted, to prove a generalization to locally compact groups of the negative result in [5]; namely, if

$$\lim_{\gamma \to \mathrm{Id}} \frac{\left\| \mathrm{Id} - T_\mu^\gamma \right\|}{1 - \left\| T_\mu^\gamma \right\|} = \infty,$$

then $T_\mu^n f$ fails to converge a.e. for some $f \in L_2(X, \beta, m)$.

Of course, one possible application of the above has already been mentioned in Section 2 with the special case of abelian groups. Another application would be that it strengthens 3.13 and 3.15 where the hypotheses had the effect of bounding the denominator away from 0. The abelian examples show that this is too strong. Theorem 3.11 can also be seen as a special case of 3.17 in some situations.

However, we also have seen that even if μ is strictly aperiodic, it can happen that $\|T_\mu^\gamma\| = 1$ for some $\gamma \neq \mathrm{Id}$ and so the hypotheses above cannot be fulfilled. Yet, if G is nilpotent, then we have observed that this is not a problem and this condition is possible, depending on the behavior of $\|T_\mu^\gamma\|$ for γ near Id. What exact centering criterion on μ for nilpotent groups is needed for the hypothesis of Theorem 3.17 to hold is not yet clear to us.

The difference between 3.11 and 3.17 is illustrated by groups like the one in Example 3.10. There no condition on μ can keep $\|T_\mu^\gamma\| < 1$ and allow the condition of 3.17 to even be a possible criterion. However, if μ is symmetric and strictly aperiodic, then 3.11 will apply. The distinction arises here because of the different types of spectral theorems that are being used. While the direct integral method of spectral analysis of T_μ is very useful throughout this article, with examples like 3.10, it is not as flexible. Then the spectral analysis of T_μ, or some more intrinsic measure-theoretic method, must be employed.

Finally, let us observe that there can also be a problem with the continuity of $\|T_\mu^\gamma\|$ as γ approaches Id. This is discussed in Dixmier [10], p. 76, where it is noted that generally $\|T_\mu^\gamma\|$ is continuous in γ when $\hat{G}$ is Hausdorff in the Fell topology, something that may fail to happen even in nilpotent groups. This is what makes it difficult to formulate subsequence principles for T_μ in general groups.

For example, suppose we want to find the behavior of

$$D^\gamma = \left\| \frac{1}{N} \sum_{k=1}^{N} \left(T_\mu^\gamma\right)^k - \left(T_\mu^\gamma\right)^N \right\|$$

for γ approaching Id, where the value of D^γ is 0. If $\gamma \mapsto \|T_\mu^\gamma\|$ is generally continuous, then we can say $D^\gamma \to 0$ as $\gamma \to \mathrm{Id}$. This is the only missing ingredient in the following subsequence theorem.

3.19. THEOREM. *Suppose the adapted probability measure $\mu \in M(G)$ has $\|T_\mu^\gamma\| \leq 1 - \delta_k$ for some $\delta_K > 0$, for all $\gamma \notin K$, K some compact neighborhood of* Id. *Suppose also that $\hat{G}$ is Hausdorff in the Fell topology. Let (X, β, m) be a probability space. Then there is a subsequence (n_s) such that for all $f \in L_2(X, \beta, m)$, $\lim_{s\to\infty} T_\mu^{n_s} f(x)$ exist a.e. and is $P_I f$.*

Proof. The proof is similar to the ones in [3], [4]. Consider

$$D_N^\gamma = \frac{1}{N}\sum_{k=1}^{N}(T_\mu^\gamma)^k - (T_\mu^\gamma)^N.$$

By continuity at Id and the control of $\|T_\mu^\gamma\|$ away from Id, there is a subsequence (n_s) such that for all $\gamma \in \hat{G}$

$$\sum_{s=1}^{\infty}\|D_{n_s}^\gamma\|_2^2 \le 5.$$

But then,

$$\begin{aligned}\sum_{s=1}^{\infty}\left\|\frac{1}{n_s}\sum_{k=1}^{n_s}T_\mu^k f - T_\mu^{n_s}f\right\|_2^2 &= \sum_{s=1}^{\infty}\int_\Lambda \|D_{n_s}^\gamma f_\lambda\|^2\,dF(\lambda)\\ &\le 5\int_\Lambda \|f_\gamma\|^2\,dF(\lambda) = 5\|f\|_2^2.\end{aligned}$$

That is,

$$\lim_{s\to\infty}\frac{1}{n_s}\sum_{h=1}^{n_s}T_\mu^k f(\lambda) - T^{n_s}f(x) = 0 \text{ a.e. for all } f \in L_2(X,\beta,m).$$

Since μ is adapted, by Theorem 3.1, we have our conclusion. □

3.20. *Remark.* E. Kaniuth has pointed out that the condition that $\hat{G}$ be Hausdorff is very restrictive. For example, if G is connected, then $\hat{G}$ is Hausdorff if and only if G contains a compact normal subgroup K such that G/K is abelian. See Baggett and Sund [1]. While if G is discrete, then $\hat{G}$ is Hausdorff if and only if the center of G is of finite index in G. See Thoma [36] for the main ingredient of this fact.

A number of interesting issues remain to be clarified in order to understand fully what can be achieved by spectral methods in general groups, in studying the a.e. and norm convergence of T_μ^n. Hopefully, the previous material gives some idea of what these issues are in general.

4. When m is Infinite

Some of the results for T_μ^n on $L_p(X,\beta,m)$ which were previously proved for probability spaces (X,β,m) extend to σ-finite measure spaces. However, there are some additional technical details. The most obvious issue is that when the action by $\{T_g\colon g \in G\}$ is ergodic and $m(X) = \infty$, then 0 is the only G-invariant function in $L_p(X, \beta, m)$, $1 \leq p < \infty$, and $T_\mu^n f$ will most often converge to 0 if there is a good result on L_p-norm convergence. Nevertheless, some discussion of the G-invariant functions in $L_p(X, \beta, m)$ is needed. See Jacobs [18] and Tempelman [34] for theorems as are given below.

Let $I_p = \{f \in L_p(X, \beta, m)\colon f$ is G-invariant$\}$, $1 \leq p \leq \infty$. For $p = \infty$, I_p will at least always contain the constants. For p, $1 \leq p \leq \infty$, let q be the index conjugate to p, $1 \leq q \leq \infty$. Let $I_p^\perp = \{f \in L_q(X, \beta, m)\colon \langle f, h\rangle = 0$ for all $h \in I_p\}$. So $I_\infty^\perp \subset \{f \in L_1(X, \beta, m)\colon \int f\,dm = 0\}$, and, if G is ergodic, then $I_p^\perp = L_q(X, \beta, m)$ for $1 \leq p < \infty$.

4.1. THEOREM. *For $1 < p < \infty$, $L_p(X, \beta, m)$ is the direct sum of I_p and $I_q^\perp$.*

The case $p = 1$, and $p = \infty$, as above are generally not true. For example, for $p = \infty$, we could be considering an ergodic action on an infinite measure space (X, β, m). Then $I_1 = \{0\}$, $I_1^\perp = L_\infty$ and I_∞ consists of the constants.

If $p = 1$, then if the action is ergodic too, $I_1 = \{0\}$ and

$$I_\infty^\perp = \left\{f \in L_1(X, \beta, m)\colon \int f\,dm = 0\right\}.$$

So $I_1 + I_\infty^\perp$ is a proper closed subspace in $L_1(X, \beta, m)$. For this reason, in the sequel, convergence theorems on $L_\infty(X, \beta, m)$ are not considered and convergence theorems in $L_1(X, \beta, m)$ are only considered on $I_\infty^\perp = \mathrm{cl}_{L_1}\{T_g f - f\colon g \in G,\ f \in L_1(X, \beta, m)\}$.

The following describes the norm convergence result that holds for infinite measure spaces (X, β, m) in the abelian case.

4.2. THEOREM. *Let G be an abelian group and suppose T is (weakly) continuous. Suppose μ is strictly aperiodic. Then*:

(a) *If $1 < p < \infty$, $T_\mu^n f$ converges in L_p-norm to the canonical projection on I_p given by the direct sum decomposition $L_p(X, \beta, m) = I_p + I_q^\perp$ for all $f \in L_p(X, \beta, m)$;*

(b) *$T_\mu^n f$ converges to 0 in L_1-norm for all $f \in I_\infty^\perp$.*

These results can be partially generalized to other groups in the same fashion as was done in Section 3. For example:

4.3. THEOREM. *Suppose μ is a probability measure on a group such that $\|T_\mu^\lambda\| < 1$ for all $\lambda \in \hat{G}$, $\lambda \neq$ Id. Then for any separable σ-finite measure space (X, β, m), if $1 < p < \infty$, $\lim_{n\to\infty} T_\mu^n f$ exists in L_p-norm and is the canonical projection $P_{I_p} f$, if $f \in L_p(X, \beta, m)$.*

4.4. *Remark.* It is not clear in this situation whether $\lim_{n\to\infty}\|\mu^n * f_0\|_1 = 0$ for all $f_0 \in L_1(G, \beta_G, \lambda_G)$, $\int f\, d\lambda_G = 0$. In particular, even for nilpotent groups, it seems to be unresolved whether Foguel's theorem used in 4.2 is still true.

4.5. Corollary. *Let μ be a strictly aperiodic probability measure in a nilpotent group. Then for any separable σ-finite measure space (X, β, m), if $f \in L_p(X, \beta, m)$, $1 < p < \infty$, $\lim_{n\to\infty} T_\mu^n f = P_{I_p} f$ in L_p-norm.*

A number of the theorems on pointwise convergence that were proved in Section 2 and 3, also are true for σ-finite spaces, generally when p is in the range, $1 < p < \infty$. Such results were generally stated without any finiteness hypothesis on m. One exception was Theorem 3.11. There the result is true with $1 < p < \infty$ in any σ-finite (X, β, m).

Acknowledgements. We would like to thank Professors Howe, Kaniuth, and Hebisch for their very helpful comments.

References

1. L. Baggett and T. Sund, *The Hausdorff dual problem for connected groups*, J. Functional Analysis **43** (1981), 60–68.
2. M. Bekka and E. Kaniuth, *Irreducible representations of locally compact groups that cannot be Hausdorff separated from the identity representation*. J. reine Angew. Math. **385** (1988), 203–220.
3. A. Bellow, R. Jones and J. Rosenblatt, *Almost everywhere convergence of weighted averages*, Math Ann. **293** (1992), 399–426.
4. ______, "Almost everywhere convergence of powers" in *Almost everywhere convergence*, Proceedings of the Conference on A.E. Convergence and Probability Theory, Spring, 1988, OSU, pp. 99–120, Academic Press, New York, 1989.
5. ______, "*Almost everywhere convergence of convolution powers*, preprint, 18 pages, to appear in Ergodic Theory and Dynamical Systems.
6. J.R. Blum and B. Eisenberg, *Generalized summing sequences and the mean-ergodic theorem*, Proc. Amer. Math. Soc. **42** (1974), 423–429.
7. J.R. Blum and J. Reich, *Pointwise ergodic theorems in locally compact abelian groups*, Pacific J. of Math. **103** (1982), 301–306.
8. J. Bourgain, *Almost sure convergence and bounded entropy*, Israel J. Math **62** (1988), 79–97.
9. Y. Derriennic and M. Lin, *Convergence of iterates of averages of certain operator representations and of convolution powers*, J. Funct. Anal. **85** (1989), 86–102.
10. J. Dixmier, *C^*-algebras* (a translation of Les C^*-algebras et leurs Représentations), North-Holland, New York, 1977.
11. J.M.G. Fell, *The dual spaces of C^*-algebras*, Trans. Amer. Math. Soc. **94** (1960), 365–402.
12. ______, *Weak containment and induced representations of groups*, Canadian J. Math **14** (1962), 237–268.
13. ______, *Weak containment and induced representations of groups II*, Trans. Amer. Math. Soc. **110** (1964), 424–447.
14. S. Foguel, *Iterates of a convolution on a non-abelian group*, Ann. Inst. Henri Poincaré **11** (1975), 199–202.
15. ______, *On iterates of convolutions*, Proc. Amer. Math. Soc. **47** (1975), 368–370.

16. S. Glasner, *On Choquet-Deny measures*, Ann. Inst. Henri Poincarè **12** (1976), 1–10.
17. E. Hewitt and K. Ross, *Abstract harmonic analysis*, Vol. I., Springer-Verlag, New York, 1963.
18. K. Jacobs, *Neuere Methoden and Ergebnisse der Ergodentheorie*, Springer-Verlag, New York, 1960.
19. V. Kaimanovich and A. Vershik, *Random walks on discrete groups: boundary and entropy*, Ann. Prob. **11** (1983), 457–490.
20. M. Lin, *Ergodic properties of an operator obtained from a continuous representation*, Ann. Inst. Henri Poincaré, **13** (1977), 321–331.
21. D. Lind, *Locally compact measure preserving flows*, Advances in Math. **15** (1975), 175–193.
22. G. Mackey, *The theory of unitary group representations*, The University of Chicago Press, Chicago, 1976.
23. P. Milnes and A. Paterson, *Ergodic sequences and a subspace of $B(G)$*, Rocky Mountain J. Math. **18** (1988) 681–694.
24. R.T. Moore,, *Measurable, continuous and smooth vectors for semigroups and group representations*, Mem. Amer. Math. Soc. No. 78, 1968.
25. V.I. Oseledec, *Markov chains, skew products and ergodic theorems for "general" dynamical systems*, Teor. veroyatnost. i Primenen. **10** (1965), 551–557.
26. J. Rosenblatt, *Uniqueness of invariant means for measure-preserving transformations*, Trans. Amer. Math. Soc. **265** (1981), 623–636.
27. ______, *Ergodic and mixing random walks on locally compact groups*, Math. Ann. **257** (1981), 31–42.
28. ______, *Ergodic group actions*, Arch. Math. **47** (1986), 263–269.
29. ______, "Universally bad sequences in ergodic theory" in *Almost everywhere convergence*, II, Proceedings of the Conference on A.E. Convergence and Probability Theory, Fall, 1989, Northwestern University, pp. 227–245, Academic Press, New York, 1991.
30. W. Rudin, *Real and complex analysis*, McGraw-Hill, New York, 1987.
31. ______, *Measure algebras on abelian groups*, Bulletin Amer. Math. Soc. **65** (1959), 227–247.
32. E. Stein, *On the maximal ergodic theorem*, Proc. Nat. Acad. Sci. **47** (1961), 1894–1897.
33. A.A. Tempelman, *Ergodic theorems for general dynamical systems*, Dokl. Akad. Nauk SSSR **176** (1967), 790–793 (English translation: Soviet Math Doklady **8** (1967), 1213–1216).
34. ______, *Ergodic theorems on groups*, Mokslas, Vilnius, 1986.
35. ______, *Ergodic Theorems for Group Actions*, Kluwer Academic Publishers, New York, 1992.
36. E. Thoma, *Über unitäre Darstellungen abzähbarer diskreter Gruppen*. Math. Ann. **153** (1964), 111–138.
37. J.H. Williamson, *A theorem on algebras of measures on topological groups*, Proc. Edinburgh Edin. Math. Soc. (2) **11** (1959), 195–206.
38. H. Yoshizawa, *On unitary representations of the free group*, Osaka Math J. **3** (1951), 55–63.

DePaul University
Chicago, Illinois
Ohio State University
Columbus, Ohio
Pennsylvania State University
University Park, Pennsylvania