J. KSIAM **Vol.5**, No.1, 25-33, 2001

ERGODICITY AND RANDOM WALKS ON A COMPACT GROUP

GEON HO CHOE

Abstract. Let G be a finite group with a probability measure. We investigate the random walks on G in terms of ergodicity of the associated skew product transformation.

1. Introduction

Random shuffling of n cards may be regarded as a random walk on the symmetric group on n symbols. The speed of convergence to the perfect random shuffle has been studied using the Fourier analysis on the symmetric group S_n of permutations on n symbols, i.e., the group representations of S_n. When a probability distribution μ on S_n is given, the probability of $g \in S_n$ is the probability of shuffling n cards according to the rule given by g. The random walk given by a finite sequence $g_1 g_2 \cdots g_k$ is the shuffling of n cards obtained by consecutively applying k shuffling methods g_i. We may say that a finitely many applications of shuffles produce perfect randomness if the probability of finding the walker at any point in S_n is $1/|S_n|$. The probability distribution for the location of the walker after k random applications of elements in S_n is given by the convolution μ^{*k}, hence the problem is to check the speed of the convergence of μ^{*k} to the Haar measure. See [1]. For additional information, see [3].

In this paper we investigate the properties of random walks on a general finite group from the viewpoint of ergodic theory, especially in terms of skew product transformations. Instead of considering the convolution of measures as done in other literature, we will focus on density functions, which are L^2-functions, hence our analysis will be simpler.

Let G be a compact group with the right-invariant Haar measure m. We consider a measurable function $h(x)$ defined on G satisfying (i) $h(x) \geq 0$, (ii) $\int_G h d\mu = 1$, (iii) h is bounded. Put

$$\text{supp } h = \{x \in G : h(x) > 0\}.$$

The function h may be regarded as a probability distribution on G: Put $d\nu = h(x)dm$. Then ν is a probability measure satisfying $\nu(\text{supp } h) = 1$. If $h = 1$ then $\nu = m$, and it is sometimes called the uniform distribution on G.

1991 *Mathematics Subject Classification.* Primary 94A17, secondary 60J10.

Key words and phrases. Random walk, finite group, group representation, skew product transformation.

We say that B is included in A modulo measure zero sets if $m(B\backslash A) = 0$ and that A and B are equal modulo measure zero sets if $m(A\backslash B) + m(B\backslash A) = 0$. If there is no danger of confusion, all set inclusions are understood as inclusions modulo measure zero sets. For finite groups there is no need to pay extra attention to such subtleties.

Recall that for $f_1, f_2 \in L^1(G, m)$ their convolution is defined by $f_1 * f_2(x) = \int_G f_1(xy^{-1})f_2(y)\,dm(y)$.

Let f^{*n} denote the n-times convolution of f with itself, i.e., $f^{*n} = \overbrace{f * f * \cdots * f}^{n}$. Observe that $f^{*n}(x) \geq 0$ and $\int_G f^{*n} dm = 1$. On a finite group G we have $f^{*n}(x) > 0$ if and only if x is a product of exactly n elements in the support of f.

2. Fourier analysis on a finite group

In this section we use results on the representations of a compact group. For the references, consult [2],[9].

Lemma 2.1. *Let (X, ν) be a probability space. (i) Let $\psi : X \to \mathbb{C}$ be a measurable function such that $|\psi(x)| \leq 1$. If $|\int_X \psi(x)d\nu| = 1$, then $\psi(x)$ is a constant of modulus 1 a.e. with respect to ν.*

(ii) Let $u : X \to \mathbb{C}^d$ be a vector-valued measurable function such that $||u(x)|| \leq 1$. If $\left\|\int_X u(x)d\nu\right\| = 1$, then $u(x)$ is a constant vector of norm 1 a.e. with respect to ν.

Proof. (i) Suppose $|\int_X \psi(x)\,d\nu| = 1$. Recall the Cauchy-Schwarz inequality $|\int fg d\nu|^2 \leq \int |f|^2 d\nu \int |g|^2 d\nu$. By taking $f = \psi$, $g = 1$, we have

$$1 = \left|\int_X fg\,d\nu\right|^2 \leq \int_X |f|^2 d\nu \int |g|^2 d\nu = \int_X |f|^2 d\nu \leq 1.$$

Since the equality holds, $f = \psi$ is a constant multiple of $g = 1$.

(ii) Suppose $|| \int_X u(x)\,d\nu|| = 1$. Put $v = \int_X u(x)d\nu$. Then $||v|| \leq 1$ and

$$1 = \left\|\int_X u(x)\,d\nu\right\|^2 = \left(\int_X u(x)\,d\nu, v\right) = \int_X (u(x), v)\,d\nu.$$

Now put $\psi(x) = (u(x), v)$ and apply the part (i). Then $(u(x), v) = \lambda$, $|\lambda| = 1$, which is possible only if $u(x) = \lambda v$. Incidentally, $\lambda = 1$ since $v = \int u(x)d\nu$. □

Definition 2.2. Given a group G and its subset S, let $\text{gen}(S)$ denote the set of all products of elements in S. (Taking inverses of elements in S are not allowed.) We say that $\text{gen}(S)$ is *generated* by S.

For a finite group G, $\text{gen}(S)$ is a subgroup. To see why, observe that for every $s \in S$ the elements $s, s^2, s^3, \ldots, s^n, \ldots$ are contained in a finite set G. So $s^j = s^k$ for some $j > k \geq 1$, and $s^{j-k} = e \in \text{gen}(S)$, hence $s^{-1} = s^{j-k-1} \in \text{gen}(S)$. Therefore, if $s_{i_1}^{p_1} \cdots s_{i_m}^{p_m}$ is in $\text{gen}(S)$ then its inverse $s_{i_m}^{-1} \cdots s_{i_1}^{-1}$ is also in $\text{gen}(S)$.

In general, $\text{gen}(S)$ need not be a subgroup if G is not finite. For example, if $G = \{z \in \mathbb{C} : |z| = 1\}$ and $S = \{e^{2\pi i\theta}\}$, θ irrational, then $\text{gen}(S) = \{e^{2\pi in\theta} : n \geq 1\}$, which does not contain the identity 1.

Given a compact group G and a representation ρ of G, $x \in G$, we define the operator norm of $\rho(x)$ by

$$||\rho(x)|| = \sup_{\vec{v} \neq \vec{0}} \frac{||\rho(x)\vec{v}||_2}{||\vec{v}||_2},$$

for the Euclidean norm $|| \cdot ||_2$ on $\mathbb{C}^{d_\rho}$. Define $\widehat{f}(\rho) = \int_G f(x)\rho(x)\, d\mu(x)$. If ρ is unitary, then $||\widehat{f}(\rho)|| \leq 1$ by the Minkowski inequality. Note that $\widehat{f_1 * f_2}(\rho) = \widehat{f_1}(\rho)\widehat{f_2}(\rho)$ and $\widehat{f^{*n}}(\rho) = \widehat{f}(\rho)^n$. For $\rho = 1$ we have $\widehat{f}(1) = 1$. Notation: (i) Let d_ρ denote the dimension of ρ, i.e., $\rho(x)$ is a $d_\rho \times d_\rho$ matrix. (ii) The set of all irreducible unitary representations of G is denoted by $\widehat{G}$.

Lemma 2.3. *Let ρ be a unitary representation of (G, m). (i) If $\widehat{f}(\rho)$ is an identity matrix for $\rho \neq 1$, then* supp f *is contained in a closed normal subgroup H of G, $H \neq G$.*

(ii) If $||\widehat{f}(\rho)|| = 1$ for $\rho \neq 1$, then supp f *is contained in a coset of a closed subgroup H of G, $H \neq G$.*

Proof. (i) If $\widehat{f}(\rho) = \int_G f(x)\rho(x)\, dm = I$, then $\rho(x) = I$ on supp f, hence supp f is included in $H = \{x : \rho(x) = I\}$.

(ii) If $|| \int_G f(x)\rho(x)\, dm|| = 1$, then there exists a constant unitary matrix A such that $\rho(x) = A$ a.e. on supp f, hence supp $f \subset \{x : \rho(x) = A\}$. Choose $g \in$ supp f such that $\rho(g) = A$. Then supp f is included in the coset gH where $H = \{x : \rho(x) = I\}$. □

Theorem 2.4. *Let (G, m) be a compact group.*

(i) If supp f *is not contained in any closed normal subgroup H of G, $H \neq G$, then $\dfrac{1}{n}\sum_{k=1}^{n} f^{*k}$ converges to 1 in L^2 as $n \to \infty$. In this case, $m(\bigcup_{k=1}^{\infty}$ supp $f^{*k}) = 1$.*

(ii) If supp f *is not contained in a coset of any closed subgroup H of G, $H \neq G$, then f^{*n} converges to 1 in L^2 as $n \to \infty$. In this case, $m($supp $f^{*n}) = 1$ for sufficiently large n and*

$$\|f^{*n}(x) - 1\|_{L^2}^2 = \sum_{\rho \neq 1} d_\rho \mathrm{Tr}\left(\widehat{f}(\rho)^n (\widehat{f}(\rho)^n)^*\right).$$

Proof. Recall that the functions $\sqrt{d_\rho}\rho_{ij}$, $\rho \in \widehat{G}$, form an orthonormal basis for $L^2(G, \mu)$. Part (ii) is simpler, and will be proved first. (ii) Lemma 2.3(ii) implies $||\widehat{f}(\rho)|| < 1$ for $\rho \neq 1$. ¿From the Fourier inversion formula, we have

$$f^{*n}(x) = \sum_{\rho \in \widehat{G}} \sqrt{d_\rho}\, \mathrm{Tr}\left(\widehat{f}(\rho)^n \sqrt{d_\rho}\, \rho(x^{-1})\right),$$

hence the Parseval's relation gives

$$\|f^{*n}(x) - 1\|_{L^2}^2 = \sum_{\rho \neq 1} d_\rho \left\| \mathrm{Tr}\left(\widehat{f}(\rho)^n \sqrt{d_\rho}\, \rho(x^{-1})\right)\right\|_{L^2}^2.$$

Since $\mathrm{Tr}(AB) = \sum_i (AB)_{ii} = \sum_{ij} A_{ij}B_{ji}$, we have

$$\mathrm{Tr}\left(\widehat{f}(\rho)^n \sqrt{d_\rho}\, \rho(x^{-1})\right) = \sum_{ij} (\widehat{f}(\rho)^n)_{ij} \sqrt{d_\rho}\, \overline{\rho(x)_{ij}},$$

and by the orthogonality relation we obtain

$$\left\| \mathrm{Tr}\left(\widehat{f}(\rho)^n \sqrt{d_\rho}\, \rho(x^{-1}) \right) \right\|_{L^2}^2 = \sum_{ij} |(\widehat{f}(\rho)^n)_{ij}|^2 = \mathrm{Tr}[\widehat{f}(\rho)^n (\widehat{f}(\rho)^n)^*],$$

which converges to zero as $n \to \infty$ since $||\widehat{f}(\rho)^n(\widehat{f}(\rho)^n)^*|| \le ||\widehat{f}(\rho)||^{2n}$ and $||\widehat{f}(\rho)|| < 1$.

(i) Lemma 2.3(i) implies $\widehat{f}(\rho) \neq I$ for $\rho \neq 1$. Put

$$M_n = \frac{1}{n} \sum_{k=1}^{n} \widehat{f}(\rho)^k = \frac{1}{n} \left[I - \widehat{f}(\rho)^{n+1} \right] \left[I - \widehat{f}(\rho) \right]^{-1}.$$

Then $||M_n|| \le \frac{1}{n} \cdot C$ for some constant $C > 0$. ¿From the Parseval's relation

$$\begin{aligned} \left\| \frac{1}{n} \sum_{k=1}^{n} f^{*k} - 1 \right\|_{L^2}^2 &= \sum_{\rho \neq 1} d_\rho \left\| \mathrm{Tr}\left(M_n \sqrt{d_\rho}\, \rho(x^{-1}) \right) \right\|_{L^2}^2 \\ &= \sum_{\rho \neq 1} d_\rho \sum_{ij} |(M_n)_{ij}|^2 \\ &= \sum_{\rho \neq 1} d_\rho \mathrm{Tr}\left(M_n M_n^* \right), \end{aligned}$$

which converges to zero as $n \to \infty$ since $||M_n M_n^*|| \le ||M_n||^2 \le C^2/n^2$. □

Lemma 2.5. *Let $\rho \neq 1$ be an irreducible unitary representation of G. If* supp f *is not contained in any closed normal subgroup H of G, $H \neq G$, then $I - \widehat{f}(\rho)$ is invertible.*

Proof. Suppose not. Then there exists a nonzero vector v such that $(I - \widehat{f}(\rho))v = 0$. Hence $\widehat{f}(\rho)v = v$ and $\int_G f(x)\rho(x)v\, dm = v$. Since $||\rho(x)v|| = ||v||$ for all $x \in G$, we have $\rho(x)v = v$ for a.e. x in supp f. ¿From Theorem 2.4(i) we observe that $\bigcup_{k=1}^{\infty} \mathrm{supp}\ f^{*k}$ is a dense subset in G. Since $\rho : G \to \mathbb{C}^{d_\rho \times d_\rho}$ is continuous, we see that $\rho(x)v = v$ for all $x \in G$, which contradicts the irreducibility of ρ. □

For a finite group G, L^2-convergence can be replaced by pointwise convergence since the L^2-norm satisfies $\|h\|^2 = \frac{1}{|G|} \sum_{x \in G} |h(x)|^2$. Or we may observe that $L^p(G)$, $1 \le p \le \infty$, is isomorphic to $\mathbb{C}^n$ where $n = |G|$, the number of elements in G and that on a finite-dimensional space all the norms are equivalent.

Consider a probability distribution on G, i.e., a nonnegative function f on G satisfying $\frac{1}{|G|} \sum_{x \in G} f(x) = 1$. For two arbitrary elements $x, y \in G$, suppose that the transition probability of moving from x to y in one step is given by $\frac{1}{|G|} f(yx^{-1})$. For example, $\frac{1}{|G|} f(y)$ is the probability of moving to y from the group identity element e. Thus we have a Markov chain given by random walks on G. A simple example is given by the 1-dimensional random walk on $G = \mathbb{Z}_n = \{0, 1, \ldots, n-1\}$ where $f(1) = \frac{n}{2} = f(-1)$. The random walker moves up or down, with probability $\frac{1}{2}$ in each case, depending on the outcome of a fair coin tossing.

In [1] the probability measure $Q(\{x\}) = \frac{1}{|G|}f(x)$ is used. In this case the L^1-norm of $Q^{*n} - U$ is considered where $U(\{x\}) = \frac{1}{|G|}$ is the Haar measure on G even though the L^2-norm is easier to compute. The reason for this is the following: If $|G|$ is even and if Q is uniformly distributed on half the points (that is, equal to $\frac{2}{|G|}$) and zero on the half. Then $||Q - U||_2 = \frac{1}{\sqrt{|G|}}$ is close to zero for $|G|$ large. This makes it difficult to compare the convergence rates of Q^{*n} to the uniformity as $n \to \infty$ for different values of $|G|$. But if we consider f instead of Q, this problem disappears: In the previous example, $|f(x) - 1| = 1$ for every x, hence $||f - 1||_2 = 1$. Furthermore, we can easily extend our argument to groups with infinitely many elements.

Let P be the transition matrix associated with the random walk, i.e., $P_{x,y} = \frac{1}{|G|}f(yx^{-1})$. Then $(P^2)_{x,z} = \sum_y P_{x,y}P_{y,z}$ is the probability to arrive at z from x in two steps, and so on. Note that

$$\begin{aligned}(f * f)(x) &= \frac{1}{|G|}\sum_{y\in G} f(xy^{-1})f(y)\\ &= |G|\sum_{y\in G} P_{y,x}P_{e,y}\\ &= |G|(P^2)_{e,x},\end{aligned}$$

and similarly $\frac{1}{|G|}f^{*k}(x) = (P^k)_{e,x}$ is the probability that the random walker is found at x at time k if he starts from e. Observe that $f^{*k}(x) > 0$ if and only if x is a product of k elements $g_1, \ldots, g_k$ such that $f(g_i) > 0$.

Theorem 2.6. *Let G be a finite group. (i) If* supp f *is not contained in any normal subgroup H of G, $H \neq G$, then every $g \in G$ is a product of elements in* supp f.

(ii) If supp f *is not contained in a coset of any subgroup H of G, $H \neq G$, then there exists N such that every $g \in G$ is a product of exactly n elements in* supp f *for any $n \geq N$.*

Proof. (i) For any small $\epsilon > 0$, Theorem 2.4(i) implies that for sufficiently large n such that $|\frac{1}{n}\sum_{k=1}^n f^{*k} - 1| < \epsilon$ for every x, hence $\frac{1}{n}\sum_{k=1}^n f^{*k} > 1 - \epsilon$, which implies $\bigcup_{k=1}^n \text{supp } f^{*k} = G$. Thus for every $g \in G$ there exists k such that $f^{*k}(g) > 0$.

(ii) For any small $\epsilon > 0$, Theorem 2.4(ii) implies that there exists N such that if $n \geq N$ then $|f^{*n} - 1| < \epsilon$ for every x, hence $f^{*n}(x) > 1 - \epsilon$. □

Remark 2.7. (i) The following statements are equivalent: (a) Markov chain is ergodic, (b) P is irreducible, and (c) supp f is not contained in any normal subgroup H of G, $H \neq G$. (ii) The following statements are equivalent: (a) Markov chain is weak-mixing, (b) Markov chain is mixing, (c) P is aperiodic, and (d) supp f is not contained in a coset of any subgroup H of G, $H \neq G$. The proofs are found in [7],[10].

When the group is the circle group and the random walk is the rotation by $e^{2\pi i\theta}$ or $e^{-2\pi i\theta}$, θ irrational, depending on the outcome of coin tossing. This was studied in [8].

3. Skew product transformation

Let G be a compact group with its right Haar measure λ, and (X, μ) a probability space and $T : X \to X$ an ergodic measure preserving transformation. Given a function $\phi : X \to G$, define a skew product transformation $T_\phi : G \times X \to G \times X$ by $(g, x) \mapsto (g \cdot \phi(x), Tx)$. Then T_ϕ preserves the product measure $\lambda \times \mu$. See [4]. The ergodicity of T_ϕ can be checked by the decomposition of $L^2(G \times X)$. It is known that the matrix coefficients of the irreducible unitary representations form an orthogonal basis for $L^2(G, \lambda)$. Take any irreducible unitary representation ρ and let (ρ_{ij}) be its matrix representation. Then

$$U_{T_\phi}(\rho_{ij}(g)f(x)) = \rho_{ij}(g\phi(x))f(Tx) = \sum_k \rho_{ik}(g)\rho_{kj}(\phi(x))f(Tx).$$

Hence we have the following U_{T_ϕ}-invariant orthogonal decomposition:

$$L^2(G \times X) = \oplus L^2_\rho(G \times X)$$

where the subspace $L^2_\rho(G \times X)$ is spanned by functions of the form $\rho_{ij}(g)f(x)$, $f \in L^2(X)$.

For $\rho = 1$ two Hilbert spaces $L^2_\rho(G, X)$ and $L^2(X)$ are identical. Let U_1 be the operator restricted on $L^2_\rho(G, X)$. Then two operators U_1 on $L^2_\rho(G, X)$ and U_T on $L^2(X)$ are unitarily equivalent. Since T is ergodic, there is no nonconstant eigenfunction of U_1 for the eigenvalue 1. The following is known but its proof is included here for the sake of completeness and notational convenience. Consult [6] for more details.

Fact 3.1. *(i) The skew product transformation $T_\phi : G \times X \to G \times X$ is not ergodic if and only if there exists an irreducible representation $\rho \neq 1$ satisfying $\rho(\phi(x))h(Tx) = h(x)$ for some $h = (h_i)_{1 \le i \le d}$, $h_i \in L^2(X)$, $h_i \neq 0$, where d is the dimension of ρ.*

(ii) It is not weak-mixing if and only if there exists an irreducible representation $\rho \neq 1$ and some constant $\lambda \in \mathbb{C}$, $|\lambda| = 1$, satisfying $\rho(\phi(x))h(Tx) = \lambda h(x)$ for some $h = (h_i)_{1 \le i \le d}$, $h_i \in L^2(X)$, $h_i \neq 0$, $f \neq 0$, where d is the dimension of ρ.

Proof. (i) Suppose T_ϕ is not ergodic. Then there exists a nonconstant function $h(g, x)$ in $L^2_\rho(G \times X)$, $\rho \neq 1$, such that $h(g\phi(x), Tx) = h(g, x)$. Put $h(g, x) = \sum_{i,j} \rho_{ij}(g)f_{ij}(x)$

and let ρ^T denote the transpose of ρ. Then

$$\begin{aligned}
h(g\phi(x), Tx) &= \sum_{i,j} \rho_{ij}(g\phi(x)) f_{ij}(Tx) \\
&= \sum_{i,j} \Big[\rho(g)\rho(\phi(x))\Big]_{ij} f_{ij}(Tx) \\
&= \sum_{i,j} \left(\sum_k \rho_{ik}(g)\rho_{kj}(\phi(x)) \right) f_{ij}(Tx) \\
&= \sum_{i,k} \rho_{ik}(g) \left(\sum_j \rho_{kj}(\phi(x)) f_{ij}(Tx) \right) \\
&= \sum_{i,j} \rho_{ij}(g) \left(\sum_k \rho_{jk}(\phi(x)) f_{ik}(Tx) \right),
\end{aligned}$$

and for every i, j we have

$$f_{ij}(x) = \sum_k \rho_{jk}(\phi(x)) f_{ik}(Tx) = \sum_k f_{ik}(Tx) \rho^T_{kj}(\phi(x)).$$

Define a matrix-valued function F on X by $F(x) = \big[f_{ij}(x)\big]$. Then $F(x) = F(Tx)\rho^T(\phi(x))$, and $F(x)^T = \rho(\phi(x))F(Tx)^T$. Choose a nonzero column of $F(x)^T$ and call it $f(x)$.

Conversely, suppose that there exists an irreducible representation $\rho \neq 1$ and a nonzero vector-valued function $f(x)$ such that $f(x) = \rho(\phi(x))f(Tx)$. Let $v(g, x)$ be a vector-valued function on $G \times X$ defined by $v(g, x) = \rho(g)f(x)$. It is not a constant function on $G \times X$ since $f(x) \neq 0$. Then

$$\begin{aligned}
v(g\phi(x), Tx) &= \rho(g\phi(x)) f(Tx) \\
&= \rho(g)\rho(\phi(x)) f(Tx) \\
&= \rho(g) f(x) = v(g, x).
\end{aligned}$$

Note that every component function of $v(g, x)$ is T_ϕ-invariant and that not all of them are constant. Therefore T_ϕ is not ergodic.

(ii) The proof is almost identical with the case (i). □

Here is a connection between random walks and skew product transformations. Let G be a finite group with a probability distribution μ such that the probability of random walk from $g \in G$ to $h \in G$ is given by $\mu(hg^{-1})$. Let $\{g_1, g_2, \ldots, g_k\} \subset G$ be the support of μ and put $p_i = \mu(g_i)$. Take an alphabet $\mathcal{A} = \{1, 2, \ldots, k\}$ of k symbols and define a Bernoulli shift space $X = \prod_1^\infty \mathcal{A}$ with the product measure defined by the probability distribution $(p_1, \ldots, p_k)$. Define the left shift transformation T by $(Tx)_i = x_{i+1}$ for $x \in X$. Recall that T is a measure preserving ergodic transformation. Define $\phi : X \to G$ by $\phi(x) = g_{x_1}$ for $x = (x_1, x_2, x_3, \ldots)$. Note that ϕ depends only on the first component in x so we write $\phi(x_1)$ to emphasize the fact. Using such ϕ we finally define the skew product transformation $T_\phi : G \times X \to G \times X$ by $T_\phi(g, x) = (g \cdot \phi(x), Tx)$. ¿From the bottom row of the following commutative diagram we see that the random walk under consideration is the projection of the skew product transformation where the projection $\pi : G \times X \to G$ is defined by $\pi(g, x) = g$.

$$\begin{tikzcd}
(g, x) \arrow[r, "{T_\phi}^n"] \arrow[d, "\pi"'] & (g \cdot \phi(x) \cdot \phi(Tx) \cdots \phi(T^{n-1}x), T^n x) \arrow[d, "\pi"'] \\
g \arrow[r, "\text{random walks}"] & g \cdot g_{x_1} \cdot g_{x_2} \cdots g_{x_n}
\end{tikzcd}$$

In the following we consider an absolutely continuous measure $f\,d\mu$. If a discrete measure ν is considered as in irrational random walks on the unit circle, we have to replace the conditions of supp f by the weak convergence of $\frac{1}{n}\sum_{k=1}^{n} \nu^{*k}$ and ν^{*n}, respectively, to obtain the ergodicity and the weak-mixing property.

Theorem 3.2. *(i) If* supp f *is not contained in any closed normal subgroup H of G, $H \neq G$, then T_ϕ is ergodic.*

(ii) If supp f *is not contained in a coset of any closed normal subgroup H of G, $H \neq G$, then T_ϕ is weak-mixing.*

Proof. We prove the case (i). The proof for the second case is similar. Suppose T_ϕ is not ergodic. Then the coboundary condition holds: There exists an irreducible unitary representation $\rho \neq 1$ and a nonzero vector-valued function $q = (q_i)_{1 \leq i \leq d}$ such that every $q_i : X \to \mathbb{C}$ is an L^2-function where d is the dimension of ρ and $\rho(\phi(x))q(Tx) = q(x)$. For the sake of notational convenience in the remainder of the proof we write $\phi(x_1)$ in place of $\phi(x)$. Then

$$\rho(\phi(x_1))q(x_2, x_3, x_4, \ldots) = q(x_1, x_2, x_3, \ldots). \tag{$*$}$$

By applying $\rho(\phi(x_1))$ to the both sides, we have

$$\begin{aligned}\rho(\phi(x_1))\rho(\phi(x_2))q(x_3, x_4, x_5, \ldots) &= \rho(\phi(x_1))q(x_2, x_3, x_4, \ldots)\\ &= q(x_1, x_2, x_3, \ldots).\end{aligned}$$

In general,

$$\rho(\phi(x_1) \cdots \phi(x_n))q(x_{n+1}, x_{n+2}, x_{n+3}, \ldots) = q(x_1, x_2, x_3, \ldots). \tag{$**$}$$

Let $[x_1, \ldots, x_n]$ denote the cylinder set $\{y \in X : y_1 = x_1, \ldots, y_n = x_n\}$ and define $u_n : \prod_1^n \{1, 2, \ldots, k\} \to \mathbb{C}^d$ by

$$u_n(x_1, \ldots, x_n) = \int_{[x_1, \ldots, x_n]} q(x)$$

where the integration of the vector-valued function q is done with respect to the remaining variables x_{n+1}, x_{n+2}, $\ldots$. If we integrate (**) on $[x_1, \ldots, x_n]$ with respect to the remaining variables x_{n+1}, x_{n+2}, $\ldots$, then

$$\rho(\phi(x_1) \cdots \phi(x_n))\, v_n = u_n(x_1, \ldots, x_n)$$

where v_n is obtained by integrating $u_n(x_1,\ldots,x_n)$ with respect to the variables $x_1,\ldots,x_n$. In other words, v_n is a convex linear combination of $u_n(x_1,\ldots,x_n)$ over $(x_1,\ldots,x_n)$. Since ρ is unitary, for every $(x_1,\ldots,x_n)$ we have

$$||u_n(x_1,\ldots,x_n)||_{\mathbb{C}^d} = ||v_n||_{\mathbb{C}^d},$$

which is possible only if $u_n(x_1,\ldots,x_n)$ is a constant vector that is nothing but v_n. Therefore for any fixed n the average of $q(x)$ on every cylinder set $[x_1,\ldots,x_n]$ is the same, hence q is a constant vector. Now we have $\rho(\phi(x_1))q = q$ from (*). Since supp f generates almost every element in G, we observe that $\rho(g)q = q$ for every $g \in G$, which in turn implies $q = \vec{0}$ by the irreducibility of ρ, and we have a contradiction. □

Remark 3.3. Kloss[5] considered the convergence of convolutions of a sequence of probability distributions on a compact group.

References

[1] P. Diaconis, *Group Representations in Probability and Statistics*, Inst. Math. Stat., Hayward, 1988.

[2] R.E. Edwards, *Integration and Harmonic Analysis on Compact Groups*, London Math. Soc. Lecture Note Series, vol. 8, Cambridge Univ. Press, Cambridge, 1972.

[3] L. Flatto, A.M. Odlyzko and D.B. Wales, *Random shuffles and group representations*, Ann. Prob. 13 (1985), 154–178.

[4] H. Furstenberg, *Strict ergodicity and transformation of the torus*, Amer. J. Math. 83 (1961), 573–601.

[5] B.M. Kloss, *Probability distributions on bicompact topological groups(English translation)*, Theor. Probability Appl. 4 (1959), 237–270.

[6] W. Parry, *Compact abelian group extensions of discrete dynamical systems*, Zeit. Wahrs. Verw. Geb. 13 (1969), 95–113.

[7] K. Petersen, *Ergodic Theory*, Cambridge Univ. Press, Cambridge, 1983.

[8] F. Su, *Convergence of random walks on the circle generated by an irrational rotation*, Trans. Amer. Math. Soc. 350 (1998), 3717–3741.

[9] A. Terras, *Fourier Analysis on Finite Groups and Applications*, London Math. Soc. Student Texts, vol. 43, Cambridge Univ. Press, Cambridge, 1999.

[10] P. Walters, *An Introduction to Ergodic Theory, 2nd ed.*, Springer-Verlag, New York, 1982.

Department of Mathematics,
Korea Advanced Institute of Science and Technology,
Taejon, Korea
choe@euclid.kaist.ac.kr