RANDOM WALKS AND APPROXIMATE INTEGRATION ON COMPACT HOMOGENEOUS SPACES

MICHAEL TAYLOR

ABSTRACT. We discuss methods of producing random walks on a compact homogeneous space X and examine how they lead to approximate evaluation of integrals of elements of various function spaces, including L^p spaces, L^p -Sobolev spaces, and Hölder spaces.

1. Introduction

We explore methods of choosing points x_j (randomly) on a compact homogeneous space X such that, given an integrable function f on X, we have

(1.1)
$$\int_{X} f \, d\mu = \lim_{N \to \infty} \frac{1}{N} \sum_{j=0}^{N-1} f(x_j),$$

at least with high probability. We assume X has a Riemannian metric, with associated volume measure μ , normalized so that $\mu(X) = 1$, and we assume that a compact Lie group G acts transitively on X, as a group of isometries.

To help put this work in perspective, we recall one approach to (1.1). Namely, let $\xi = (x_j)$ be a sequence of points chosen independently and randomly on X, with probability distribution μ . That is to say,

(1.2)
$$\xi \in \prod_{j=1}^{\infty} (X, \mu) = (\Xi, \nu),$$

Received November 10, 2012; received in final form March 5, 2013. 2010 Mathematics Subject Classification. 43A85, 47A35, 60G50.

the infinite Cartesian product, with product measure. Write the right-hand side of (1.1) as $\lim_{N\to\infty} \mathcal{A}_N(\xi)f$, where

(1.3)
$$A_N(\xi)f = \frac{1}{N} \sum_{i=0}^{N-1} f(x_i).$$

Clearly

$$(1.4) f \equiv 1 \implies \mathcal{A}_N(\xi)f \equiv 1.$$

Also, if

(1.5)
$$\int_{X} f \, d\mu = 0, \qquad f \in L^{2}(X, \mu),$$

then independence implies orthogonality, and we have

(1.6)
$$\|\mathcal{A}_N(\cdot)f\|_{L^2(\Xi,\nu)}^2 = \frac{1}{N^2} \sum_{i=0}^{N-1} \|f\|_{L^2(X,\mu)}^2 = \frac{1}{N} \|f\|_{L^2(X,\mu)}^2,$$

which tends to 0 as $N \to \infty$ like C/N, so

(1.7)
$$\|\mathcal{A}_N(\cdot)f\|_{L^2(\Xi,\nu)} = \|f\|_{L^2(X,\mu)}^2 N^{-1/2}.$$

Putting together (1.4) and (1.5)–(1.7), we have

(1.8)
$$f \in L^{2}(X,\mu), \qquad \int_{X} f \, d\mu = a$$
$$\implies \|\mathcal{A}_{N}(\cdot)f - a\|_{L^{2}(\Xi,\nu)} = \|f - a\|_{L^{2}(X,\mu)} N^{-1/2},$$

which by Chebyshev's inequality implies, for $\kappa > 0$,

(1.9)
$$\nu\left(\left\{\xi \in \Xi : \left| \mathcal{A}_{N}(\xi)f - a \right| \ge \frac{\kappa}{\sqrt{N}} \right\}\right) \le \kappa^{-2} \|f - a\|_{L^{2}(X,\mu)}^{2}.$$

This quantifies the statement that, with high probability, one can approximate $\int_X f \, d\mu$ to within $CN^{-1/2}$ by (1.3). This result is a special case of the weak law of large numbers.

We recall that the strong law of large numbers provides results on convergence

(1.10)
$$\mathcal{A}_N(\xi)f \longrightarrow \int_X f \, d\mu,$$

for ν -a.e. $\xi \in \Xi$. One approach to this is to consider the one-sided shift

(1.11)
$$\sigma: \Xi \longrightarrow \Xi, \qquad \sigma(x_1, x_2, x_3, \dots) = (x_2, x_3, x_4, \dots).$$

Given $f \in L^p(X, \mu), 1 \le p < \infty$, define

(1.12)
$$F \in L^p(\Xi, \nu), \qquad F(\xi) = f(x_1).$$

Then

(1.13)
$$\mathcal{A}_N(\xi)f = \frac{1}{N} \sum_{i=0}^{N-1} F(\sigma^i(\xi)),$$

and the result

(1.14)
$$\mathcal{A}_N(\xi)f \longrightarrow \int_X f \, d\mu = \int_{\Xi} F \, d\nu, \quad \nu\text{-a.e., and in } L^p(\Xi, \nu),$$

follows from Birkhoff's Ergodic theorem, given that

(1.15)
$$\sigma$$
 in (1.11) is ergodic,

which is equivalent to the statement that

(1.16)
$$G \in L^2(\Xi, \nu)$$
, $G = G \circ \sigma$ ν -a.e. \Longrightarrow $G = \text{const.}$, a.e.

We mention that (1.15) is a special case of Proposition 3.1 below, though one can give a proof of (1.15) that is much simpler than our proof of Proposition 3.1. (Cf. [15], p. 201.)

While the result (1.8)–(1.9) is called "weak" and the result (1.14) is called "strong," we note that the former result is not strictly weaker than the latter, since the former is accompanied by an explicit quantitative estimate.

Implementing the approximation to $\int_X f \, d\mu$ described above requires one to have a method for picking $x_j \in X$ randomly, with probability distribution μ . We mention some cases where such random choices are available. One case is the circle $S^1 = \mathbb{R}/(2\pi\mathbb{Z})$, with measure given by $d\theta/2\pi$, which is measure theoretically equivalent to the interval I = [0,1]. Many high level programming languages have canned routines to pick random points here. Taking products yields such random choices on tori \mathbb{T}^n . One can also choose random points on the unit sphere $S^{n-1} \subset \mathbb{R}^n$, as follows. Using a change of variable, one can take the uniform distribution dt on $\{t \in [0,1]\}$ to the Gaussian distribution $\pi^{-1/2}e^{-t^2}dt$ on \mathbb{R} , and, taking products, choose random points $x \in \mathbb{R}^n$ with Gaussian distribution $\pi^{-n/2}e^{-|x|^2}dx$. Then the standard projection $\mathbb{R}^n \setminus 0 \to S^{n-1}$ given by $x \mapsto x/|x|$ gives random points on S^{n-1} . Consequently, we can produce random points on $SU(2) \approx S^3$, hence on SO(3), which is covered by SU(2), and also on SO(4), which is covered by $SU(2) \times SU(2)$.

Moving on to other cases, such as $X = G = SO(n), n \ge 5$, other SO(n)-homogeneous spaces, such as Grassmannians, and other compact Lie groups and associated homogeneous spaces, it is not so easy to pick random points, and we are motivated to produce other methods to carry out (1.1).

Here we discuss a method that involves random walks on G, and associated random walks on X. There is a substantial literature on random walks on groups, including [8], [7], [6], [3], [12], [4], [5], [10], [13], [1], [9], [2], [16], and references given there. However, the particular problems we treat here are somewhat different from the problems these works emphasize. Here, we study how uniformly distributed on X are individual paths of random walks. The

emphasis on the papers mentioned above is on how close to uniform is the probability distribution μ_k of the location, at step k, of all possible random walks. Of course, there are points of contact, particularly involving analysis of the operator A, defined below in (1.29). We give more details on this at the end of Section 2.

To illustrate the distinction, suppose one picks independent, random paths ξ_j , fixes k, and takes x_j to be the position of ξ_j at step k. Then the right side of (1.1) tends, not to $\int_X f \, d\mu$, but to $\int_X f \, d\mu_k$. Instead, one might evaluate the jth random path ξ at step k_j , with $k_j \nearrow \infty$. In either case, one throws away the bulk of the intermediate steps of the random walks, and this is not what we want to do.

We continue with a description of the program of this paper. Let Y be a compact Hausdorff space, equipped with a probability measure P (frequently, Y will be a finite point set), and let

$$(1.17) R: Y \longrightarrow G$$

be continuous. Given $\omega_j \in Y$, we write $R_{\omega_j} \in G$. We assume the following.

(1.18) We have a method to pick random points in Y,

with distribution given by P. Form

$$(1.19) \qquad (\Omega, Q) = \prod_{k=1}^{\infty} (Y, P),$$

with product measure Q, and denote a point in Ω by $\omega = (\omega_1, \omega_2, \omega_3, ...)$. Given $f \in L^p(X, \mu)$, set

$$(1.20) T(\omega_1)f(x) = f(R_{\omega_1}x),$$

and, for $k \in \mathbb{N}$,

(1.21)
$$T_k(\omega)f(x) = T(\omega_1)\cdots T(\omega_k)f(x)$$
$$= f(R_{\omega_k}\cdots R_{\omega_1}x).$$

Then set

(1.22)
$$\mathcal{A}_N(\omega)f(x) = \frac{1}{N} \sum_{k=0}^{N-1} T_k(\omega)f(x).$$

We will show that analogues of (1.8)–(1.9) and (1.14) hold, that is,

(1.23)
$$\mathcal{A}_N(\omega)f \longrightarrow \int_X f \, d\mu,$$

either with high probability, or for Q-a.e. $\omega \in \Omega$, under certain natural conditions, essentially (modulo some natural technicalities) that

(1.24) the group generated by
$$g_0^{-1}R(Y) \subset G$$
 is dense in G ,

where we fix $y_0 \in Y$ and set $g_0 = R_{y_0}$. In Section 2 we produce analogues of (1.8)–(1.9) and in Section 3 we produce analogues of (1.14). Both constructions involve variants of the arguments sketched above, though they are necessarily substantially more elaborate, requiring further tools, including techniques from representation theory and harmonic analysis.

We describe two types of examples of particular interest.

EXAMPLE 1.1. G = SO(n). We take the points x_j in (1.1) to be defined in terms of products of randomly chosen reflections. In more detail, we set

$$(1.25) Y = S^{n-1} \times S^{n-1}, \omega_j = (\alpha_j, \beta_j), \alpha_j, \beta_j \in S^{n-1},$$

and

$$(1.26) R_{\omega_j} = R_{\alpha_j} R_{\beta_j},$$

where, given $\alpha \in S^{n-1}$, R_{α} is the reflection across the (n-1)-plane orthogonal to α . Construction of random points in S^{n-1} was described above. The fact that each $g \in SO(n)$ is a product of (an even number of) reflections (n of them if n is even, n-1 if n is odd) implies (1.24).

EXAMPLE 1.2. G is a general compact Lie group, and $Y = \{g_0, \ldots, g_K\} \subset G$ is a finite subset, with $R: Y \hookrightarrow G$ the inclusion. We pick $p_j \in (0,1)$ such that $\sum_{j=0}^{K} p_j = 1$, and assign probability p_j to g_j . We make the hypothesis (1.24).

In the following sections, we see how (1.23) works in the context of (1.24), with particular emphasis on Examples 1.1–1.2. To describe the results, set

(1.27)
$$g = f + \overline{g}, \quad \overline{g} = \int_X g \, d\mu, \int_X f \, d\mu = 0.$$

In Section 2 we show that, for $g \in L^2(X)$,

(1.28)
$$\|\mathcal{A}_{N}(\cdot)g - \overline{g}\|_{L^{2}(X \times \Omega)}^{2}$$

$$\leq \frac{1}{N} \left\{ \|f\|_{L^{2}(X)} + 2 \sum_{k=1}^{N-1} \|A^{k}f\|_{L^{2}(X)} \right\} \|f\|_{L^{2}(X)},$$

where

(1.29)
$$A = \int_{Y} T(\omega_1) dP(\omega_1).$$

The right-hand side of (1.28) vanishes as $N \to \infty$ provided $A^k \to 0$ strongly on

$$L_0^2(X) = \left\{ f \in L^2(X) : \ \int_X f \, d\mu = 0 \right\},$$

that is,

If $A^k \to 0$ in norm on $L_0^2(X)$, which implies

(1.31)
$$||A^k f||_{L^2(X)} \le Be^{-\alpha k} ||f||_{L^2(X)}, \quad \forall f \in L^2_0(X),$$

we get

(1.32)
$$\|\mathcal{A}_{N}(\cdot)g - \overline{g}\|_{L^{2}(X \times \Omega)} \leq \frac{C}{N^{1/2}} \|f\|_{L^{2}(X)},$$

parallel to (1.7). We give general conditions guaranteeing (1.30) or (1.31), and see that for Example 1.1, (1.31) holds, while for Example 1.2, the hypothesis (1.24) implies (1.30).

In Section 3 we show that, if $f \in L^p(X)$, $1 \le p < \infty$,

(1.33)
$$\mathcal{A}_N(\omega)f(x) \longrightarrow \int_X f d\mu, \quad \mu \times Q\text{-a.e., and in } L^p(X \times \Omega),$$

as $N \to \infty$, provided the following ergodic property holds:

(1.34)
$$g \in L^2(X), \quad T(\omega_1)g = g \text{ for } P\text{-a.e. } \omega_1 \in Y$$
 $\Longrightarrow g \text{ constant}, \mu\text{-a.e.}$

This result is parallel to the strong law of large numbers. It does not have the quantitative features of (1.28) and (1.32).

In Section 4, we interpolate the L^2 estimates of Section 2 and other estimates to estimate $\|\mathcal{A}_N(\cdot)g - \overline{g}\|_{L^p(X \times \Omega)}$, for $1 . Using the fact that the operators <math>\mathcal{A}_N(\omega)$ on $L^p(X)$ commute with powers of the Laplace operator, we also estimate

(1.35)
$$\int_{\Omega} \| \mathcal{A}_{N}(\omega) g - \overline{g} \|_{H^{s,p}(X)}^{p} dQ(\omega),$$

where $H^{s,p}(X)$ denotes the L^p -Sobolev space of regularity index s. Via the Sobolev embedding theorem, this leads, under the hypothesis (1.31), to

$$(1.36) \qquad \int_{\Omega} \|\mathcal{A}_{N}(\omega)g - \overline{g}\|_{C(X)}^{p} dQ(\omega) \leq CN^{-p/p^{\#}} \|g - \overline{g}\|_{H^{s,p}(X)}^{p},$$

valid for $sp > \dim X$, where

(1.37)
$$\frac{p}{p^{\#}} = 1, \quad 2 \le p < \infty,$$

$$p - 1, \quad 1$$

As a corollary of this, we produce an estimate of the left side of (1.36) when q satisfies a Lipschitz-Hölder condition.

NOTE. For notational convenience, we often write $L^p(X)$ and $L^p(X \times \Omega)$ in place of $L^p(X, \mu)$ and $L^p(X \times \Omega, \mu \times Q)$.

2. Quantitative L^2 results

We take notation as in Section 1 and assume $f \in L^2(X, \mu)$. Clearly

$$(2.1) f \equiv 1 \implies \mathcal{A}_N(\omega)f \equiv 1.$$

We next take

(2.2)
$$f \in L_0^2(X, \mu)$$
, i.e., $\int_X f \, d\mu = 0$.

We have

the last identity because each $T_k(\omega)$ and $T_\ell(\omega)$ are unitary on $L^2(X)$. To proceed, set

(2.4)
$$A = \int_{Y} T(\omega_1) dP(\omega_1).$$

We have (for $\ell > k$)

(2.5)
$$\int_{\Omega} T(\omega_{k+1}) \cdots T(\omega_{\ell}) dQ(\omega) = A^{\ell-k}.$$

If we take f to be real valued and note that $T(\omega_j)$ is reality-preserving, we obtain

(2.6)
$$\|\mathcal{A}_{N}(\cdot)f\|_{L^{2}(X\times\Omega)}^{2}$$

$$= \frac{1}{N^{2}} \left\{ N\|f\|_{L^{2}(X)}^{2} + 2 \sum_{\ell \leq k \leq N-1} (A^{k-\ell}f, f)_{L^{2}(X)} \right\}$$

$$\begin{split} &= \frac{1}{N^2} \bigg\{ N \|f\|^2 \|_{L^2(X)} + 2 \sum_{k=1}^{N-1} (N-k) \big(A^k f, f\big)_{L^2(X)} \bigg\} \\ &= \frac{1}{N} \bigg\{ \|f\|_{L^2(X)}^2 + 2 \sum_{k=1}^{N-1} \bigg(1 - \frac{k}{N}\bigg) \big(A^k f, f\big)_{L^2(X)} \bigg\}. \end{split}$$

As we will see below, an appropriately precise version of (1.24) implies

$$(2.7) \forall f \in L_0^2(X), ||A^k f||_{L^2(X)} \le \varepsilon_k(f) \to 0 as k \to \infty.$$

This leads to the following result.

PROPOSITION 2.1. When (2.7) holds, then, for $f \in L_0^2(X)$,

(2.8)
$$\lim_{N \to \infty} \| \mathcal{A}_N(\cdot) f \|_{L^2(X \times \Omega)} = 0.$$

Proof. The identity (2.6) plus (2.7) gives

(2.9)
$$\|\mathcal{A}_N(\cdot)f\|_{L^2(X\times\Omega)}^2 \le \frac{1}{N} \left\{ \|f\|_{L^2(X)} + 2\sum_{k=1}^{N-1} \varepsilon_k(f) \right\} \|f\|_{L^2(X)},$$

which clearly yields (2.8).

Below we will also see cases where the following stronger version of (2.7) holds, namely there exist $B, \alpha \in (0, \infty)$ such that

$$(2.10) \forall f \in L_0^2(X), ||A^k f||_{L^2(X)} \le Be^{-\alpha k} ||f||_{L^2(X)}.$$

This leads to the following result.

PROPOSITION 2.2. When (2.10) holds, there exists $C \in (0, \infty)$ such that, for all $f \in L_0^2(X)$,

(2.11)
$$\|\mathcal{A}_N(\cdot)f\|_{L^2(X\times\Omega)} \le C\|f\|_{L^2(X)}N^{-1/2}.$$

Proof. This follows from (2.6) together with the estimate

(2.12)
$$\sum_{k=1}^{N-1} \left| \left(A^k f, f \right)_{L^2(X)} \right| \le B \sum_{k=1}^{N-1} e^{-\alpha k} \|f\|_{L^2(X)}^2$$
$$\le \frac{B'}{\alpha} e^{-\alpha} \|f\|_{L^2(X)}^2,$$

which readily yields (2.11).

Putting together (2.1) and the results of Propositions 2.1–2.2, we have the following.

Proposition 2.3. Take $g \in L^2(X)$ and set

(2.13)
$$\overline{g} = \int_X g \, d\mu.$$

If (2.7) holds, then

where $\delta_N(g-\overline{g}) \to 0$ as $N \to \infty$, and if (2.10) holds, then

(2.15)
$$\|\mathcal{A}_{N}(\cdot)g - \overline{g}\|_{L^{2}(X \times \Omega)} \le C\|g - \overline{g}\|_{L^{2}(X)}N^{-1/2}.$$

Note. Given (2.13),
$$\|g - \overline{g}\|_{L^2(X)}^2 = \|g\|_{L^2(X)}^2 - \overline{g}^2 \le \|g\|_{L^2(X)}^2$$
.

We now look at conditions on $R: Y \to G$ under which (2.7) or (2.10) hold. For this, it is convenient to make an orthogonal direct sum decomposition

(2.16)
$$L_0^2(X,\mu) = \bigoplus_{\ell > 1} V_{\ell},$$

where each V_{ℓ} is a finite dimensional subspace of $L_0^2(X)$ on which G acts irreducibly; denote the action $\pi_{\ell}: G \to U(V_{\ell})$. Possibly dim $V_{\ell} = 1$ for some ℓ , but each π_{ℓ} is nontrivial. Clearly each operator $T(\omega_1)$ leaves each space V_{ℓ} invariant, and hence so does A, given by (2.4). Set

$$(2.17) A_{\ell} = A|_{V_{\ell}}.$$

Note that $||A_{\ell}|| \leq 1$. The following is immediate.

PROPOSITION 2.4. The result (2.7) holds if and only if, for each $\ell \geq 1$, there exists $k = k(\ell)$ such that

$$||A_{\ell}^{k(\ell)}|| < 1.$$

The result (2.10) holds provided there exists $K \ge 1$ and $\theta < 1$ such that

$$||A_{\ell}^{K}|| \le \theta < 1, \quad \forall \ell \ge 1.$$

To proceed, note that

(2.20)
$$A_{\ell} = \pi_{\ell}(\lambda) = \int_{C} \pi_{\ell}(g) \, d\lambda(g),$$

where λ is the probability measure on G given by pushing P forward:

$$(2.21) \lambda = R_*(P),$$

i.e.,

(2.22)
$$\int_{G} u(g) d\lambda(g) = \int_{Y} u(R_{\omega_1}) dP(\omega_1).$$

Note that λ is supported on $R(G) \subset G$. Also, for $k \in \mathbb{N}$,

$$(2.23) A_{\ell}^{k} = \pi_{\ell}(\lambda^{(k)}),$$

where $\lambda^{(k)}$ is the k-fold convolution product:

(2.24)
$$\lambda^{(k)} = \lambda * \cdots * \lambda \quad (k \text{ factors}).$$

We also have

(2.25)
$$\lambda^{(k)} = R_*^{(k)} (P^{(k)}),$$

where

$$(2.26) R^{(k)}: Y \times \cdots \times Y \longrightarrow G, R^{(k)}(\omega_1, \dots, \omega_k) = R_{\omega_k} \cdots R_{\omega_k},$$

and $P^{(k)} = P \times \cdots \times P$ is the associated product probability measure on $Y^k = Y \times \cdots \times Y$.

We are now ready to show that cases covered in Example 1.2 satisfy hypothesis (2.7).

PROPOSITION 2.5. Let $Y = \{g_0, g_1, \dots, g_K\} \subset G$ and $P(\{g_j\}) = p_j \in (0, 1)$ for each j. Assume the group generated by $g_0^{-1}Y = \{e, h_1, \dots, h_K\}$ is dense in G. Then (2.7) holds.

Proof. We will show that, for all $\ell \geq 1$,

which, as noted in Proposition 2.4, implies (2.7). Since V_{ℓ} is finite dimensional, it suffices to show

$$(2.28) ||A_{\ell}v|| < ||v||, \text{for all nonzero } v \in V_{\ell}.$$

In the present case, we have $\lambda = \sum p_j \delta_{g_j}$, hence

(2.29)
$$A_{\ell}v = \pi_{\ell}(\lambda)v = \sum_{j>0} p_j \pi_{\ell}(g_j)v,$$

so

(2.30)
$$\pi_{\ell}(g_0)^{-1}A_{\ell}v = p_0v + \sum_{j>1} p_j\pi_{\ell}(h_j)v.$$

The only way to avoid (2.28) is for

whenever $h \in g_0^{-1}Y$. Then (2.31) must hold for all h in the group generated by $g_0^{-1}Y$, hence, if this is dense, for all $h \in G$. Since π_{ℓ} is irreducible, this implies $V_{\ell} = \operatorname{Span}(v)$ and π_{ℓ} acts trivially on V_{ℓ} . However, as noted above, a trivial representation of G does not occur in the decomposition (2.16). This gives (2.28) and completes the proof.

We now extend Proposition 2.5 to the more general setting, involving $R: Y \to G$ and $\lambda = R_*(P)$. Define $\widetilde{Y} \subset R(Y)$ by

$$(2.32) \widetilde{Y} = \{g \in G : \lambda(B_r(g)) > 0 \ \forall r > 0\},\$$

where $B_r(g)$ is the ball of radius r centered at g. The proof of Proposition 2.5 readily extends, to establish the following.

PROPOSITION 2.6. Take \widetilde{Y} as in (2.32) and $g_0 \in \widetilde{Y}$. If the group generated by $g_0^{-1}\widetilde{Y}$ is dense in G, then (2.28) holds, hence (2.7) holds.

REMARK. Proposition 2.6 can also be proven using arguments as in the proof of Proposition 10 (p. 68) of [7].

We next show that the stronger condition (2.10) holds for Example 1.1. The key to this is to note that if k is large enough that each element of SO(n) is a product of 2k reflections, then the k-fold convolution product $\lambda^{(k)}$, given by (2.24) and (2.25), is absolutely continuous with respect to Haar measure on G. We have the following result.

PROPOSITION 2.7. With $R: Y \to G$ and $\lambda = R_*(P)$, as in (2.21), suppose there exists $k \in \mathbb{N}$ such that

(2.33)
$$\lambda^{(k)} = \psi_k \, dg, \quad \psi_k \in L^1(G).$$

Then there exists $K \in \mathbb{N}$ such that (2.19) holds, hence (2.10) holds.

Proof. In this case, Proposition 2.6 applies, and we have

What we need is uniformity, which we can get from (2.33), in view of the following result, which is the Riemann–Lebesgue lemma for compact Lie groups.

LEMMA 2.8. If
$$\psi \in L^1(G)$$
 and $\pi_{\ell}(\psi) = \int_G \pi_{\ell}(g)\psi(g) dg$, then (2.35) $\|\pi_{\ell}(\psi)\| \longrightarrow 0$, as $\ell \to \infty$.

Proof. This lemma follows by approximating ψ in L^1 -norm by an element of $L^2(G)$ and using the group Plancherel theorem. In more detail, given $\psi \in L^1(G), \varepsilon > 0$, find $\varphi \in L^2(G)$ such that $\|\psi - \varphi\|_{L^1} < \varepsilon$. The Plancherel theorem gives

(2.36)
$$\|\varphi\|_{L^{2}(G)}^{2} = \sum_{\ell} \|\pi_{\ell}(\varphi)\|_{HS}^{2}$$

$$\geq \sum_{\ell} \|\pi_{\ell}(\varphi)\|^{2},$$

where $\|\cdot\|_{\mathrm{HS}}$ denotes the Hilbert–Schmidt norm. Hence

(2.37)
$$\|\pi_{\ell}(\varphi)\| \longrightarrow 0 \text{ as } \ell \to \infty,$$

and so

(2.38)
$$\limsup_{\ell \to \infty} \|\pi_{\ell}(\psi)\| \le \|\psi - \varphi\|_{L^{1}(G)} + \limsup_{\ell \to \infty} \|\pi_{\ell}(\varphi)\| < \varepsilon,$$

giving
$$(2.35)$$
.

As mentioned in the Introduction, analyses of the operator A, given by (2.4), play a role in many of the papers on random walks cited there. We have presented some simple results on A in Propositions 2.5–2.7, applicable to our estimates on how the right side of (1.1) approaches the left side. We mention further results that others have obtained.

In the setting of Example 1.1, very interesting and detailed results on the behavior of $\lambda^{(k)}$, appearing in (2.23)–(2.25), are given in [10]. In particular, there is a fairly narrow interval of ks, centered around $(n/2) \log n$, over which there is a transition from $\|\psi_k - 1\|_{L^1} \approx 2$ to $\|\psi_k - 1\|_{L^1} \ll 1$ (for ψ_k as in (2.33)). This "cutoff phenomenon" mirrors that discovered in the context of the symmetric groups in [6]. Note, however, that such a cutoff phenomenon does not enhance the rate of convergence given in (2.11).

There are versions of Propositions 2.5–2.6 with conclusions more precise than (2.27). We describe a result given in Theorem 6 of the recent paper [16]. As shown there, in the setting of Proposition 2.6, if G is a compact, connected, semisimple Lie group, there exist $A = A(G) \in (0, \infty)$ and $c = c(\lambda) \in (0, \infty)$ such that

where $-\gamma_{\ell}$ is the eigenvalue of Δ (the Laplace–Beltrami operator on X) on V_{ℓ} , which is necessarily an eigenspace of Δ .

Note that (2.19) holds, hence (2.10) holds, if there exists c > 0 such that

$$(2.40) ||A_{\ell}|| \le 1 - c, \quad \forall \ell \ge 1.$$

If this holds, we say A has a "spectral gap." In the symmetric case, where $\lambda(\mathcal{O}) = \lambda(\mathcal{O}^{-1})$ for each Borel set $\mathcal{O} \subset G$, this is in fact equivalent to (2.19). Given such a symmetry condition, Proposition 2.7 provides a sufficient condition for such a spectral gap condition to hold. Note that Proposition 2.7 does not apply to probability measures λ that are finite sums of point masses. Some important spectral gap results have been obtained in this setting. The first such result was obtained in [3]. Theorem 1.1 and Corollary 1.2 of [3] imply that there exist such atomic probability measures on G with a spectral gap whenever G is a compact simple Lie group, not locally isomorphic to either SO(3) or SO(4). Going further, [1] treated the case G = SU(2), and gave explicit sufficient conditions on elements $g_j \in SU(2)$ such that the spectral gap condition (2.40) holds for

(2.41)
$$\lambda = \frac{1}{2N} \sum_{i=1}^{N} \{ \delta_{g_i} + \delta_{g_j^{-1}} \}.$$

These results have been extended to SU(n) for n > 2 in [2]. Noncommutativity plays a crucial role here. It is easy to check that such a spectral gap phenomenon for atomic measures cannot hold when G is a compact Abelian Lie group.

3. L^p and Q-a.e. convergence

We form the Cartesian product $X \times \Omega$, with product measure $\mu \times Q$, and define

$$(3.1) \varphi: X \times \Omega \longrightarrow X \times \Omega, \varphi(x, \omega_1, \omega_2, \dots) = (R_{\omega_1} x, \omega_2, \omega_3, \dots),$$

which is readily seen to be measure preserving. Given $f \in L^p(X), 1 \le p < \infty$, define

(3.2)
$$F \in L^p(X \times \Omega), \quad F(x, \omega) = f(x).$$

Then

(3.3)
$$\mathcal{A}_N(\omega)f(x) = \frac{1}{N} \sum_{j=0}^{N-1} F(\varphi^j(x,\omega)).$$

Birkhoff's Ergodic theorem implies there exists $F^{\#} \in L^p(X \times \Omega)$ such that, as $N \to \infty$,

(3.4)
$$\mathcal{A}_N(\omega)f(x) \longrightarrow F^{\#}(x,\omega), \quad \mu \times Q$$
-a.e., and in $L^p(X \times \Omega)$,

and $F^{\#}$ has the invariance property

(3.5)
$$F^{\#} \circ \varphi = F^{\#}, \quad \mu \times Q \text{-a.e.}$$

Of course, we also have

(3.6)
$$\int_{X\times\Omega} F^{\#} d\mu dQ = \int_{X\times\Omega} F d\mu dQ = \int_{X} f d\mu.$$

We can go from here to the assertion that (1.23) holds, in $L^p(X \times \Omega)$ and $\mu \times Q$ -a.e., provided φ in (3.1) is ergodic. The following is a key result in that direction.

Proposition 3.1. Let $G \in L^2(X \times \Omega)$ satisfy

(3.7)
$$G \circ \varphi = G, \quad \mu \times \Omega \text{-a.e.}$$

Then there exists $g \in L^2(X)$ such that

(3.8)
$$G(x,\omega) = g(x), \quad \text{for } \mu \times Q\text{-a.e. } (x,\omega).$$

Furthermore,

(3.9)
$$g(R_{\omega_1}x) = g(x), \quad \text{for } \mu \times P\text{-a.e. } (x, \omega_1).$$

Proof. To begin, take an orthonormal basis $\{b_j : j \geq 0\}$ of $L^2(Y, P)$, such that $b_0 \equiv 1$. Then an orthonormal basis of $L^2(\Omega, Q)$ is given by

$$(3.10) b_{\alpha}(\omega) = b_{\alpha_1}(\omega_1)b_{\alpha_2}(\omega_2)\cdots, \quad \alpha = (\alpha_1, \alpha_2, \ldots),$$

where $\alpha_j \geq 0$ and all but finitely many α_j are = 0. Take $G \in L^2(X \times \Omega)$, satisfying (3.7), and write

(3.11)
$$G(x,\omega) = \sum_{\alpha} G_{\alpha}(x)b_{\alpha_1}(\omega_1)b_{\alpha_2}(\omega_2)\cdots.$$

Note that

$$(3.12) G(R_{\omega_1}x, \omega_2, \omega_3, \dots) = \sum_{\alpha} G_{\alpha}(R_{\omega_1}x)b_{\alpha_1}(\omega_2)b_{\alpha_2}(\omega_3) \dots$$
$$= \sum_{\alpha} G_{(\alpha_2, \alpha_3, \dots)}(R_{\omega_1}x)b_{\alpha_2}(\omega_2)b_{\alpha_3}(\omega_3) \dots,$$

so the identity (3.7) is equivalent to

(3.13)
$$\sum_{\alpha_1} G_{(\alpha_1, \alpha_2, \dots)}(x) b_{\alpha_1}(\omega_1) = G_{(\alpha_2, \alpha_3, \dots)}(R_{\omega_1} x),$$

for all $(\alpha_2, \alpha_3, \dots)$. Now note that

(3.14)
$$||G||_{L^{2}(X \times \Omega)}^{2} = \sum_{\alpha} ||G_{\alpha}||_{L^{2}(X)}^{2}.$$

Hence

(3.15)
$$||G_{(\alpha_2,\alpha_3,\dots)}||_{L^2(X)}^2 = \int_Y ||G_{(\alpha_2,\alpha_3,\dots)}(R_{\omega_1})||_{L^2(X)}^2 dP(\omega_1)$$

$$= \sum_{\alpha_1} ||G_{(\alpha_1,\alpha_2,\dots)}||_{L^2(X)}^2,$$

the last identity by (3.13). Inductively, for each $k \in \mathbb{N}$,

(3.16)
$$||G_{(\alpha_{k+1},\alpha_{k+2},\dots)}||_{L^2(X)}^2 = \sum_{\alpha_1,\dots,\alpha_k} ||G_{(\alpha_1,\alpha_2,\dots)}||_{L^2(X)}^2.$$

Given $G_{\alpha} \neq 0$, we know that there exists k such that $\alpha_j = 0$ for all $j \geq k + 1$, so (3.16) implies

$$(3.17) G_{\alpha} \neq 0 \implies \alpha = (0, 0, 0, \dots).$$

This gives (3.8), and (3.9) readily follows.

We recall that, in order to establish ergodicity, it suffices to check invariant functions in L^2 . Hence, Proposition 3.1 yields the following.

Proposition 3.2. Assume that

(3.18)
$$g \in L^2(X), \quad g(R_{\omega_1} x) = g(x) \quad \text{for } \mu \times P\text{-a.e. } (x, \omega_1) \in X \times Y$$

 $\implies g = const., \quad \mu\text{-a.e.}$

Then φ in (3.1) is ergodic. Hence, given $f \in L^p(X), 1 \leq p < \infty$,

(3.19)
$$\mathcal{A}_N(\omega)f(x) \longrightarrow \int_X f d\mu, \quad \mu \times Q\text{-a.e.}, \text{ and in } L^p(X \times \Omega),$$

as $N \to \infty$.

REMARK. Recalling $T(\omega_1)$, given by (1.20), we see that the hypothesis (3.18) is equivalent to the following:

(3.20)
$$g \in L^2(X), \quad T(\omega_1)g = g \text{ for } P\text{-a.e. } \omega_1 \in Y$$

 $\implies g \text{ is constant}, \quad \mu\text{-a.e.}$

4. Quantitative L^p and Sobolev space results

We produce further estimates on $\mathcal{A}_N(\omega)g - \overline{g}$, where

$$(4.1) \overline{g} = \int_X g \, d\mu.$$

For the sake of simplicity, we work under the hypothesis (2.10), with consequence (2.15), that is,

The interested reader could make parallel arguments under hypothesis (2.7), with consequence (2.14). To begin, we extend (4.2) to L^p -estimates, using interpolation. In fact, with

(4.3)
$$\mathfrak{A}_N g(x,\omega) = \mathcal{A}_N(\omega)g(x) - \overline{g},$$

the estimate (4.2) says

(4.4)
$$\mathfrak{A}_N: L^2(X,\mu) \longrightarrow L^2(X \times \Omega, \mu \times Q), \\ \|\mathfrak{A}_N g\|_{L^2(X \times \Omega)} \le CN^{-1/2} \|g - \overline{g}\|_{L^2(X)}.$$

On the other hand, it is clear that

(4.5)
$$\mathfrak{A}_{N}: L^{p}(X, \mu) \longrightarrow L^{p}(X \times \Omega, \mu \times Q), \\ \|\mathfrak{A}_{N}g\|_{L^{p}(X \times \Omega)} \leq \|g - \overline{g}\|_{L^{p}(X)},$$

for all $p \in [1, \infty]$, in particular for p = 1 and $p = \infty$. The Riesz–Thorin interpolation theorem gives, for 1 ,

(4.6)
$$\|\mathfrak{A}_N g\|_{L^p(X \times \Omega)} \le C N^{-1/p^{\#}} \|g - \overline{g}\|_{L^p(X)},$$

with

(4.7)
$$p^{\#} = p, \quad 2 \le p < \infty, \\ p', \quad 1 < p \le 2,$$

where p' is the dual exponent to p. In turn, we rewrite (4.6) as

for 1 . Note that the pth power of the left side of (4.8) is

(4.9)
$$\int_{\Omega} \| \mathcal{A}_{N}(\omega) g - \overline{g} \|_{L^{p}(X)}^{p} dQ(\omega).$$

Chebyshev's inequality then yields, for $\kappa > 0$,

$$(4.10) Q(\lbrace \omega \in \Omega : ||S_N(\omega)g - \overline{g}||_{L^p(X)} \ge \kappa N^{-1/p^{\#}} \rbrace)$$

$$\leq \frac{C}{\kappa^p} ||g - \overline{g}||_{L^p(X)}^p.$$

We now obtain stronger estimates, under the hypothesis that, for some $s > 0, p \in (1, \infty)$,

$$(4.11) g \in H^{s,p}(X).$$

Here $H^{s,p}$ is an L^p -Sobolev space. One way to characterize it is as follows. Let Δ denote the Laplace–Beltrami operator on X. Then

(4.12)
$$H^{s,p}(X) = (1 - \Delta)^{-s/2} L^p(X),$$

i.e., $u \in H^{s,p}(X) \Leftrightarrow (1-\Delta)^{s/2}u \in L^p(X)$. More details can be found in [14], Chapter 13. The key to the use of these spaces here is the fact that, since $x \mapsto R_{\omega_j}x$ is an isometry on X for each $\omega_j \in Y$, $(1-\Delta)^{s/2}$ commutes with $T(\omega_1)$, given by (1.20), hence

(4.13)
$$\mathcal{A}_N(\omega)(1-\Delta)^{s/2}g = (1-\Delta)^{s/2}\mathcal{A}_N(\omega)g,$$

for all $\omega \in \Omega, N \in \mathbb{N}, s \in \mathbb{R}^+, g \in H^{s,p}(X)$. We also have

$$(4.14) (1 - \Delta)^{s/2} 1 \equiv 1.$$

Hence, given $g \in H^{s,p}(X)$, we can replace g by $(1-\Delta)^{s/2}g$ in (4.8) and (4.10), apply (4.13) and (4.14), and deduce the following.

PROPOSITION 4.1. Assume that (2.10) holds. Given $g \in H^{s,p}(X)$, s > 0, $p \in (1, \infty)$, we have

$$(4.15) \qquad \int_{\Omega} \|\mathcal{A}_{N}(\omega)g - \overline{g}\|_{H^{s,p}(X)}^{p} dQ(\omega) \le CN^{-p/p^{\#}} \|g - \overline{g}\|_{H^{s,p}(X)}^{p},$$

and hence, for $\kappa > 0$,

$$(4.16) Q(\left\{\omega \in \Omega : \left\| \mathcal{A}_{N}(\omega)g - \overline{g} \right\|_{H^{s,p}(X)} \ge \kappa N^{-1/p^{\#}} \right\})$$

$$\leq \frac{C}{\kappa^{p}} \|g - \overline{g}\|_{H^{s,p}(X)}^{p}.$$

We can apply to this the Sobolev embedding result

$$(4.17) H^{s,p}(X) \subset C(X), for sp > \dim X,$$

where C(X) denotes the space of continuous functions on X, with the sup norm, to deduce the following.

COROLLARY 4.2. In the setting of Proposition 4.1, if $sp > \dim X$, then

$$(4.18) \qquad \int_{\Omega} \|\mathcal{A}_{N}(\omega)g - \overline{g}\|_{C(X)}^{p} dQ(\omega) \leq CN^{-p/p^{\#}} \|g - \overline{g}\|_{H^{s,p}(X)}^{p},$$

and there is a similar replacement for (4.16).

We draw a further corollary, using spaces $C^r(X)$, defined for $r \geq 0$ as follows. If $r = k \in \mathbb{Z}^+$, $C^r(X) = C^k(X)$ consists of functions whose derivatives of order $\leq k$ are continuous. If $r = k + \sigma$, $k \in \mathbb{Z}^+$, $0 < \sigma < 1$, $C^r(X) = C^{k+\sigma}(X)$

consists of functions whose derivatives of order $\leq k$ are Hölder continuous with exponent σ . We have

$$(4.19) C^r(X) \subset H^{s,p}(X), \quad \forall s < r, p \in (1, \infty).$$

Hence, Corollary 4.2 implies the following.

COROLLARY 4.3. In the setting of Proposition 4.1, if

$$(4.20) 0 < r \le \frac{1}{2} \dim X \quad and \quad p > \frac{\dim X}{r},$$

then

$$(4.21) \qquad \int_{\Omega} \left\| \mathcal{A}_{N}(\omega) g - \overline{g} \right\|_{C(X)}^{p} dQ(\omega) \leq C_{rp} N^{-1} \left\| g - \overline{g} \right\|_{C^{r}(X)}^{p}.$$

For $X = \mathbb{T}^n$, the *n*-dimensional torus, there are sharper estimates for a natural class of non-random integral approximations of functions satisfying Hölder conditions; see, for example, [11].

Acknowledgments. Research partially supported by NSF grants DMS-0758320 and DMS-1161620. Thanks to the referee for valuable suggestions for improvements.

References

- J. Bourgain and A. Gamburd, On the spectral gap for finitely-generated subgroups of SU(2), Invent. Math. 171 (2008), 83–121. MR 2358056
- [2] J. Bourgain and A. Gamburd, A spectral gap theorem in SU(d), J. Eur. Math. Soc. (JEMS) 14 (2012), 1455–1511. MR 2966656
- [3] C. Chou, A. Lau and J. Rosenblatt, Approximation of compact operators by spaces of translations, Illinois J. Math. 29 (1985), 340–350. MR 0784527
- [4] P. Diaconis and L. Saloff-Coste, Random walks on finite groups: A survey of analytic techniques, Probability measures on groups and related structures XI (Oberwolfach, 1994), World Scientific, River Edge, NJ, 1995, pp. 44–75. MR 1414925
- [5] P. Diaconis and L. Saloff-Coste, Walks on generating sets of groups, Invent. Math. 134 (1998), 251–299. MR 1650316
- [6] P. Diaconis and M. Shahshahani, Generating a random permutation with random transpositions, Z. Wahrsch. Verw. Gebiete 57 (1981), 159–179. MR 0626813
- [7] Y. Guivarc'h, M. Keane and B. Roynette, Marches Aléatoires sur les Groupes de Lie, Lecture Notes in Mathematics, vol. 624, Springer, New York, 1977. MR 0517359
- [8] H. Kesten, Symmetric random walks on groups, Trans. Amer. Math. Soc. 92 (1959), 336–354. MR 0109367
- [9] G. Lebeau and L. Michel, Semiclassical analysis of a random walk on a manifold, Ann. Probab. 38 (2010), 277–315. MR 2599200
- [10] U. Porod, The cut-off phenomenon for random reflections, Ann. Probab. 24 (1996), 74–96. MR 1387627
- J. Rauch and M. Taylor, Qualitative estimates for multidimensional integrals, Houston J. Math. 36 (2010), 727–749. MR 2727000
- [12] J. Rosenthal, Random rotation: Characters and random walks on SO(N), Ann. Probab. 22 (1994), 348–423. MR 1258882
- [13] L. Saloff-Coste, Probability on groups: Random walks and invariant diffusions, Notices Amer. Math. Soc. 48 (2001), 968–977. MR 1854532

- [14] M. Taylor, Partial differential equations, vol. 1–3, Springer, New York, 1996. MR 1395147
- [15] M. Taylor, Measure theory and integration, Graduate Studies in Mathematics, vol. 76, Amer. Math. Soc., Providence, RI, 2006. MR 2245472
- $[16]\ \ P.\ Varju,\ Random\ walks\ on\ compact\ groups,\ 2012,\ available\ at\ {\tt arXiv:1209.1745v2}.$

MICHAEL TAYLOR, MATHEMATICS DEPARTMENT, UNIVERSITY OF NORTH CAROLINA, CHAPEL HILL, NC 27599, USA

 $E ext{-}mail\ address: met@email.unc.edu}$