Publ. RIMS, Kyoto Univ.
29 (1993), 153-159

Remarks on the Shuffling Problem for Finite Groups

By

Akihito HORA*

Abstract

The shuffling problem is discussed as the asymptotic behavior of random walks on finite groups. We give a new characterization for asymptotic equidistribution of such random walks in terms of representations of the group. As applications, we characterize perfect groups and consider random walks on classical Weyl groups.

§1. Introduction and Main Results

A set S is shuffled by successive random actions. Taking a group G acting on S and a probability measure μ on G, we let elements of G act on S successively and independently in the frequency controlled by μ. G should act transitively on S to shuffle the whole of S. Thus a pair (G, μ) defines a shuffling rule of S.

Let us give a rigid definition in terms of random variables. Let $X_1, X_2, \cdots$ be G-valued independent random variables with the same distribution μ. Put $X_0 \equiv e$ (=the identity element of G) a.s. We call their product $W_n = X_n X_{n-1} \cdots X_1 X_0$ a left random walk on G generated by μ. The mapping $W_n : S \to S$ gives an n times shuffle of S. In this setting, shuffling rule (G, μ) is considered to 'work well' if and only if the distribution of W_n, which is equal to μ^{*n} (n-fold convolution of μ), tends to the uniform distribution on G as $n \to \infty$. We thus concern ourselves with asymptotic behavior of infinite convolution of probabilities on G. We can set up the following two questions.

Question (A) When does shuffling rule (G, μ) work well?

Question (B) Then, how many times shuffle should we repeat to reach a sufficiently shuffled condition?

In this note we give a comprehensive answer to Question (A) in the case

Communicated by T. Kawai, May 13, 1992.
1991 Mathematics Subject Classification: 60B15.

* Department of Mathematics, College of Liberal Arts and Sciences, Okayama University, Okayama 700, Japan.

where G is a finite group. Although this problem is already treated in the pioneering work due to Kawada-Ito [8] and Heyer [6], we give here a new characterization in terms of representations of G (Theorem 1). This characterization enables us to understand a relation between the extent of noncommutativity of G and the shuffling effect by G. As an application of this result, we characterize perfect groups by means of asymptotic behavior of random walks on them (Theorem 2). Question (B) is deeper than (A). Solutions are given in some concrete models, in which interesting cut-off phenomena are discovered. See Aldous-Diaconis [1] and Diaconis [3, 4].

Let us recall some terminology on Markov chains to state our theorems. We use the notations of $\boldsymbol{N}=\{0, 1, 2, \cdots\}$ and $\boldsymbol{N}_+=\boldsymbol{N}\setminus\{0\}$. Let $(W_n)_{n\in\boldsymbol{N}}$ be a left (or right) random walk on a finite group G as above and $p_n(s, t)$ $(s, t\in G, n\in\boldsymbol{N}_+)$ its nth transition probability. If every state communicates with the other states, i.e. $\forall s, t\in G$, $\exists n\in\boldsymbol{N}_+$, $p_n(s, t)>0$, $(W_n)_{n\in\boldsymbol{N}}$ is said to be irreducible. Then, since the greatest common divisor d of $\{n\in\boldsymbol{N}_+;\ p_n(t, t)>0\}$ is independent of t, d is called the period of $(W_n)_{n\in\boldsymbol{N}}$. If $(W_n)_{n\in\boldsymbol{N}}$ has no nontrivial period (i.e. if $d=1$), $(W_n)_{n\in\boldsymbol{N}}$ is said to be aperiodic. Lastly, if the distribution of W_n converges to the normalized Haar measure m of G (at every point $t\in G$), $(W_n)_{n\in\boldsymbol{N}}$ is said to be asymptotically equidistributed (after Heyer [6]). Of course this means that shuffling rule (G, μ) works well. The support of μ is denoted by $\operatorname{supp}\mu$. $\langle B\rangle$ denotes the subgroup of G generated by a subset B.

Theorem 1. *Let G be a finite group, μ a probability on G and $(W_n)_{n\in\boldsymbol{N}}$ a left (or right) random walk on G generated by μ. Then the following conditions are equivalent.*

(1) *$(W_n)_{n\in\boldsymbol{N}}$ is asymptotically equidistributed.*

(2) *$\exists k\in\boldsymbol{N}_+$ such that $(\operatorname{supp}\mu)^k=G$.*

(3) *$(W_n)_{n\in\boldsymbol{N}}$ is irreducible and aperiodic.*

(4) *$\langle\operatorname{supp}\mu\rangle=G$ holds and $\operatorname{supp}\mu$ is not contained in any coset of any proper normal subgroup of G.*

(5) *$\langle\operatorname{supp}\mu\rangle=G$ holds and no nontrivial one-dimensional characters of G are constant-valued on $\operatorname{supp}\mu$.*

Remark 1. In Section 2 we give a self-contained proof of this theorem in order of (1)$\Rightarrow$(2)$\Rightarrow$(3)$\Rightarrow$(4)$\Rightarrow$(5)$\Rightarrow$(1). Our main step is (5)$\Rightarrow$(1), where we use Fourier analysis on G, especially estimation of the spectral radius of $\hat{\mu}(\tau)(\tau\in\hat{G})$.

Remark 2. Equivalence (1)$\Leftrightarrow$(3) follows from a convergence theorem for doubly stochastic Markov chains with finite states. See e.g. Feller [5] Chapter XV or Chung [2] Section 6. Equivalence (1)$\Leftrightarrow$(4) is due to Kawada-Ito [8] and (1)$\Leftrightarrow$(2)$\Leftrightarrow$(4) is in Heyer [6] Chapter II. In many cases, however, our condition (5) is a convenient criterion for asymptotic equidistribution.

Remark 3. Condition (4) is strictly weaker than:
(4′) $\operatorname{supp}\mu$ is not contained in any coset of any proper subgroup of G.
See Proposition 1 in Section 3. It seems that (4) is sometimes confused with (4′). Even if $\operatorname{supp}\mu$ is contained in a proper coset, μ does not necessarily generate a periodic random walk.

As an application of Theorem 1, we obtain the following theorem. A finite group G is called a perfect group if $G=G'$ (the commutator subgroup of G). Perfect groups have strong noncommutativity.

Theorem 2. *G is a perfect group if and only if every irreducible random walk on G is asymptotically equidistributed.*

The proof is given in Section 2.

§2. Proofs of Theorem 1 and Theorem 2

2.1. Proof of Theorem 1
We discuss only left random walks. Notations in Section 1 are freely used.

Lemma 1. *For probabilities μ and ν on G, $\operatorname{supp}(\mu*\nu)=\operatorname{supp}\mu\ \operatorname{supp}\nu$.*

Proof. From $(\mu*\nu)(t)=\sum_{s\in G}\mu(ts^{-1})\nu(s)$, we have $(\mu*\nu)(t)>0 \Leftrightarrow \exists s\in G$ such that $\mu(ts^{-1})>0$ and $\nu(s)>0 \Leftrightarrow t=ts^{-1}s\in\operatorname{supp}\mu\ \operatorname{supp}\nu$. ■

Lemma 2. *$(W_n)_{n\in N}$ is irreducible if and only if $\langle\operatorname{supp}\mu\rangle=G$.*

Proof. (only if part): Let $t\in G$. Since $\exists n\in\boldsymbol{N}_+$ such that $0<p_n(e,t)=\sum_{s_1,\cdots,s_{n-1}\in G}p(e,s_1)p(s_1,s_2)\cdots p(s_{n-1}t)=\sum_{s_1,\cdots,s_{n-1}\in G}\mu(s_1)\mu(s_2s_1^{-1})\cdots\mu(ts_{n-1}^{-1})=\mu^{*n}(t)$, we have $t\in\operatorname{supp}\mu^{*n}=(\operatorname{supp}\mu)^n\subset\langle\operatorname{supp}\mu\rangle$ (by Lemma 1).

(if part): $\forall t\in G$, $\exists n\in\boldsymbol{N}_+$ such that $t, t^{-1}\in(\operatorname{supp}\mu)^n$. Then, $p_n(e,t)=\mu^{*n}(t)>0$. Since $p(s_1,s_2)=p(s_1s',s_2s')$ therefore $p_n(s_1,s_2)=p_n(s_1s',s_2s')$, we have $p_n(t,e)=p_n(e,t^{-1})>0$. Hence e communicates with every state t. ■

Proof of (1)⇒(2): Since $\mu^{*n}(t)\to m(t)=|G|^{-1}>0$ for $\forall t\in G$, we have $G=\operatorname{supp}\mu^{*n}=(\operatorname{supp}\mu)^n$ for sufficiently large $n\in\boldsymbol{N}_+$ by Lemma 1. ■

Proof of (2)⇒(3): Irreducibility immediately follows from Lemma 2. Let $d\in\boldsymbol{N}_+$ denote the period of $(W_n)_{n\in N}$. Since $p_k(e,e)=\mu^{*k}(e)$ for $\forall k\in\boldsymbol{N}_+$, we have $\{k\in\boldsymbol{N}_+;\ (\operatorname{supp}\mu)^k=G\}\subset\{k\in\boldsymbol{N}_+;\ p_k(e,e)>0\}\subset\boldsymbol{N}_+d$. However, $(\operatorname{supp}\mu)^k=G$ implies $G=G\ (\operatorname{supp}\mu)^{-1}(\operatorname{supp}\mu)\subset G\ \operatorname{supp}\mu=(\operatorname{supp}\mu)^k\ \operatorname{supp}\mu=(\operatorname{supp}\mu)^{k+1}$. Hence, if $\{k\in\boldsymbol{N}_+;\ (\operatorname{supp}\mu)^k=G\}\neq\emptyset$, we have $d=1$. ■

Proof of (3)⇒(4): Let $(W_n)_{n\in N}$ be irreducible. It suffices to show that

$(W_n)_{n\in N}$ is periodic if $\exists N \lhd G$, $N\neq G$ and $\exists s\neq e$ such that $\operatorname{supp}\mu\subset sN$. Let r denote the order of $sN\in G/N$ ($r\in \boldsymbol{N}_+$, $r\geqq 2$). Then, $\operatorname{supp}\mu\subset sN$, $\cdots$, $(\operatorname{supp}\mu)^{r-1}\subset s^{r-1}N$, $(\operatorname{supp}\mu)^r\subset N$ hold and sN, $\cdots$, $s^{r-1}N$, N are disjoint as subsets of G. Hence, $p_k(e, e)>0$ ($\Leftrightarrow e\in\operatorname{supp}\mu^{*n}=(\operatorname{supp}\mu)^k$ by Lemma 1) $\Rightarrow (\operatorname{supp}\mu)^k\subset N\Rightarrow r\mid k$. $(W_n)_{n\in N}$ thus has a period divided by r. ■

Proof of (4)$\Rightarrow$(5): Let χ be a nontrivial one-dimensional character of G such that $\chi(t)=\chi(t_0)$ for $\forall t\in\operatorname{supp}\mu$. Then $\operatorname{supp}\mu\subset t_0\ker\chi$ where $\ker\chi\lhd G$, $\ker\chi\neq G$. ■

Before proving (5)$\Rightarrow$(1), we recall some properties of Fourier transform on G. Let $\hat{G}$ denote the equivalence classes of irreducible representations of G. We can regard $\hat{G}$ as the family of unitary representatives.

Definition. For $\boldsymbol{C}$-valued functions f, g on G, we define

○ Fourier transform: $\hat{f}(\tau)=\sum_{t\in G} f(t)\tau(t^{-1})$ $(\tau\in\hat{G})$

○ Inner product: $(f\mid g)=\mid G\mid^{-1}\sum_{t\in G} f(t)\overline{g(t)}$.

Facts 1) $\widehat{f*g}(\tau)=\hat{g}(\tau)\hat{f}(\tau)$

2) $\hat{m}(\tau)=\delta(\tau, 1)$ where m is the normalized Haar measure of G, 1 is the trivial representation of G and δ is Kronecker's delta.

3) (Plancherel's formula)

$$(f\mid g)=\frac{1}{\mid G\mid^2}\sum_{\tau\in\hat{G}} d_\tau \operatorname{trace}(\hat{f}(\tau)\hat{g}(\tau)^*)$$

where d_τ is the degree of τ and $*$ denotes the adjoint operator.

Proof of (5)$\Rightarrow$(1) is divided into four steps. We show the $L_2(G)$-norm of $\mu^{*n}-m$ converges to 0 as $n\to\infty$.

[Step 1] Using Facts 1), 2), 3), we have

$$\|\mu^{*n}-m\|_2^2=(\mu^{*n}\mid\mu^{*n})-2(\mu^{*n}\mid m)+(m\mid m)=\frac{1}{\mid G\mid^2}\sum_{\tau\neq 1} d_\tau \operatorname{trace}(\hat{\mu}(\tau)^n\hat{\mu}(\tau)^{*n}).$$

Let $r(\cdot)$ and $\|\cdot\|$ denote the spectral radius and the operator norm respectively. Since μ is a probability and τ is unitary, we have

$$r(\hat{\mu}(\tau))\leqq\|\hat{\mu}(\tau)\|\leqq\sum_{t\in G}\mu(t)\|\tau(t^{-1})\|=1.$$

[Step 2] We show that $\operatorname{trace}(\hat{\mu}(\tau)^n\hat{\mu}(\tau)^{*n})\to 0$ as $n\to\infty$ if $r(\hat{\mu}(\tau))<1$. Since

$$\operatorname{trace}(\hat{\mu}(\tau)^n\hat{\mu}(\tau)^{*n})\leqq\|\hat{\mu}(\tau)^n\|_{HS}^2\leqq d_\tau\|\hat{\mu}(\tau)^n\|^2 \quad\text{and}\quad \lim_{n\to\infty}\|\hat{\mu}(\tau)^n\|^{1/n}=r(\hat{\mu}(\tau))<1,$$

we can take δ $(r(\hat{\mu}(\tau))<\delta<1)$ such that, for $n\in\boldsymbol{N}_+$ large enough, $\operatorname{trace}(\hat{\mu}(\tau)^n\hat{\mu}(\tau)^{*n})\leqq d_\tau\delta^{2n}(\to 0$ as $n\to\infty)$.

[Step 3] We show that $d_\tau=1$ if $r(\hat{\mu}(\tau))=1$. Let V be the representation space of τ. $r(\hat{\mu}(\tau))=1$ implies $\exists\lambda\in\boldsymbol{C}$ and $\exists v\in V$ such that $|\lambda|=1$, $\|v\|=1$ and $\hat{\mu}(\tau)v=\lambda v$. Through the isometric identification $V\simeq\boldsymbol{R}^{2d_\tau}$, we denote by B the closed unit ball of V. Since $\lambda v, \tau(t^{-1})v\in\partial B$ in the equality

$$\lambda v=\hat{\mu}(\tau)v=\sum_{t\in\operatorname{supp}\mu}\mu(t)\tau(t^{-1})v \qquad \text{(convex combination)},$$

we have $\tau(t^{-1})v=\lambda v$ for $\forall t\in\operatorname{supp}\mu$ from uniform convexity of B. Expressing $\forall t\in G$ as $t=t_1\cdots t_l$ where $t_1,\cdots,t_l\in\operatorname{supp}\mu$, we have $\tau(t^{-1})v=\tau(t_l^{-1})\cdots\tau(t_1^{-1})v=\lambda^l v$. Thus v spans an invariant subspace of V. Since τ is irreducible, $d_\tau=\dim V=1$ holds.

[Step 4] If one-dimensional character χ is not constant on $\operatorname{supp}\mu$, then $|\hat{\mu}(\chi)|=|\sum_{t\in\operatorname{supp}\mu}\mu(t)\overline{\chi(t)}|<1$. Hence, from Step 1, 2, 3, we consequently obtain

$$\|\mu^{*n}-m\|_2^2=\frac{1}{|G|^2}\sum_{\chi\neq1,\,d_\chi=1}|\hat{\mu}(\chi)|^{2n}+\frac{1}{|G|^2}\sum_{d_\tau>1}d_\tau\operatorname{trace}(\hat{\mu}(\tau)^n\hat{\mu}(\tau)^{*n})\xrightarrow{n\to\infty}0\,. \quad\blacksquare$$

Remark. Unless $\langle\operatorname{supp}\mu\rangle=G$, condition (5) gets strictly weaker than (4). In fact, we have only to take a perfect, but not simple, group G and probability μ supported by a proper normal subgroup of G.

2.2. Proof of Theorem 2

[Step 1] Let G be perfect. Since $\#\{\text{one-dimensional character of } G\}=|G:G'|=1$, G has no nontrivial one-dimensional characters. Hence, Theorem 1 (1)$\Rightarrow$(5) and Lemma 2 show that every irreducible random walk is asymptotically equidistributed.

[Step 2] Let G be not perfect. Since G/G' is a nontrivial abelian group, we have $\exists A$: abelian group and $\exists C$: nontrivial cyclic group such that $G/G'\simeq A\oplus C$. Let $\pi: G\to G/G'$ be the canonical homomorphism. Then $\pi^{-1}A\lhd G$, $\pi^{-1}A\neq G$. Since $G/\pi^{-1}A\simeq\pi G/A\simeq C$, we can take $t\in G$ such that coset $t\pi^{-1}A$ generates G. Thus, letting $\operatorname{supp}\mu=t\pi^{-1}A$, we see from Theorem 1 (1)$\Leftrightarrow$(4) and Lemma 2 that μ generates an irreducible random walk which is not asymptotically equidistributed. $\blacksquare$

§ 3. Some Results and Examples

3.1. Supplementary Results

Theorem 1 tells us the following fact.

Corollary 1. *Let G be a finite group and μ a probability on G such that $\operatorname{supp}\mu\ni e$. Then every irreducible random walk generated by μ is asymptotically equidistributed.*

To let e act is just to do nothing. It may be less efficient to do shuffling

busily without any pauses.

Next we prove the assertion mentioned at Remark 3 in Section 1.

Proposition 1. *There exist a finite group G and a subset S of G satisfying* i) $\langle S\rangle=G$, ii) *S is not contained in any coset of any proper normal subgroup of G,* iii) *S is contained in a coset of a proper subgroup of G.*

Proof. Let G be a noncommutative simple group. Since G is not a p-group, G contains a nontrivial proper subgroup H_0 by Sylow's theorem. Taking $t_1\notin H_0$, we put $H_1=\langle t_1H_0\rangle$. Next, taking $t_2\notin H_1$ if $H_1\neq G$, we put $H_2=\langle t_2H_1\rangle$. We thus inductively get $H_n=\langle t_nH_{n-1}\rangle$ and $t_n\notin H_{n-1}$. Since $|H_{n-1}|<|H_n|$ clearly holds, we have $G=\langle t_lH_{l-1}\rangle$ for some $l\in\boldsymbol{N}_+$. $S=t_lH_{l-1}$ satisfies the desired conditions. ■

3.2. Examples

We begin with a very simple case.

Example 1. $G=\boldsymbol{Z}/n\boldsymbol{Z}$ $(n\in\boldsymbol{N}_+)$. Probability μ generates an asymptotically equidistributed random walk if and only if $\forall p$: prime divisor of n, $\exists s$, $t\in\operatorname{supp}\mu$ such that $s\not\equiv t\pmod p$.

By virtue of Theorem 1 we can check asymptotic equidistribution of random walks on a group whose one-dimensional characters we know well. Let us here mention, as a natural extension of random walks caused by symmetric groups (i.e. card shuffling), random walks by the actions of classical Weyl groups. These are regarded as some random reflections in Euclidean spaces or as random walks between the Weyl chambers, while they can be interpreted as some more concrete models. We give necessary and sufficient conditions (AE) for random walks generated by μ on G to be asymptotically equidistributed, where G is either Weyl group $\mathcal{W}(A_l)$, $\mathcal{W}(B_l)(=\mathcal{W}(C_l))$ or $\mathcal{W}(D_l)$. See Humphreys [7] Chapter III for structures of Weyl groups and their actions.

Example 2. $G=\mathcal{W}(A_l)\simeq\mathfrak{S}_{l+1}$ $(l\geqq 1)$. G causes shuffles of $l+1$ cards (as they are turned down). The one-dimensional characters of G make a group $\{\chi_0,\chi_1\}$ of order 2, where $\chi_0(\sigma)=1$ and $\chi_1(\sigma)=\operatorname{sgn}\sigma$ for $\sigma\in G$. Hence our conditions (AE) are:

○ $\langle\operatorname{supp}\mu\rangle=G$

○ $\operatorname{supp}\mu$ contains both even permutations and odd ones.

Example 3. $G=\mathcal{W}(B_l)\simeq(\boldsymbol{Z}/2\boldsymbol{Z})^l\rtimes\mathfrak{S}_l$ $(l\geqq 2)$. In this case 'turning over' is added as new operations to permutations. The one-dimensional characters of G make $\{\chi_0,\chi_1,\chi_2,\chi_3\}\simeq(\boldsymbol{Z}/2\boldsymbol{Z})^2$. See Serre [9] Section 8.2 for irreducible

representations of semidirect product groups. We express an element $t\in G$ as $t=a\sigma$ where $a=(a(1), \cdots, a(l))\in(\boldsymbol{Z}/2\boldsymbol{Z})^l$ and $\sigma\in\mathfrak{S}_l$, which are regarded as a sign change and a permutation respectively. Then we put $\chi_0(t)=1$, $\chi_1(t)=\operatorname{sgn}\sigma$, $\chi_2(t)=(-1)^{a(1)+\cdots+a(l)}$ and $\chi_3(t)=\chi_1(t)\chi_2(t)$ for $t=a\sigma$. Now we put $A_\mu=\{a\in(\boldsymbol{Z}/2\boldsymbol{Z})^l$; $\exists\sigma\in\mathfrak{S}_l$ such that $a\sigma\in\operatorname{supp}\mu\}$ and $S_\mu=\{\sigma\in\mathfrak{S}_l$; $\exists a\in(\boldsymbol{Z}/2\boldsymbol{Z})^l$ such that $a\sigma\in\operatorname{supp}\mu\}$. Our conditions (AE) are:

- ○ $\langle\operatorname{supp}\mu\rangle=G$
- ○ A_μ contains both even sign changes and odd ones.
- ○ S_μ contains both even permutations and odd ones
- ○ $\operatorname{supp}\mu$ contains an element $a\sigma$ such that both a and σ are even or both are odd.

Example 4. $G=\mathcal{W}(D_l)\simeq(\boldsymbol{Z}/2\boldsymbol{Z})^{l-1}\rtimes\mathfrak{S}_l(l\geqq 4)$. We use the same expression $t=a\sigma$, $a\in(\boldsymbol{Z}/2\boldsymbol{Z})^{l-1}$, $\sigma\in\mathfrak{S}_l$ and notation S_μ as in Example 3. Similarly, we have (AE):

- ○ $\langle\operatorname{supp}\mu\rangle=G$
- ○ S_μ contains both even permutations and odd ones.

References

[1] Aldous, D. and Diaconis, P., Shuffling cards and stopping times, *Amer. Math. Monthly*, **93** (1986), 333-348.

[2] Chung, K. L., *Markov chains with stationary transition probabilities, second edition*, Berlin Heidelberg New York, Springer, 1967.

[3] Diaconis, P., Applications of non-commutative Fourier analysis to probability problems, In: Hennequin, P. L. (ed.) *École d'été de probabilités de Saint-Flour XV-XVII* 1985-1987. (Lect. Notes Math., **1362**, pp. 51-100) Berlin Heidelberg New York, Springer, 1988.

[4] Diaconis, P., *Group representations in probability and statistics*, Hayward, California, Inst. Math. Stat., 1988.

[5] Feller, W., *An introduction to probability theory and its applications, vol. 1, third edition*, New York London Sydney, John Wiley & Sons, Inc. 1968.

[6] Heyer, H., *Probability measures on locally compact groups*, Berlin Heidelberg New York, Springer, 1977.

[7] Humphreys, J. E., *Introduction to Lie algebras and representation theory*, New York Heidelberg Berlin, Springer, 1972.

[8] Kawada, Y. and Ito, K., On the probability distribution on a compact group. I, *Proc. Phys.-Math. Soc. Japan*, **22** (1940), 977-998.

[9] Serre, J.-P., *Linear representations of finite groups*, New York Heidelberg Berlin. Springer, 1977.