ON TRANSIENCE OF CARD SHUFFLING

SARA BROFFERIO AND WOLFGANG WOESS

Laboratoire de Probabilités, Université de Paris 6
4 Place Jussieu, 75252 Paris, France

Dipartimento di Matematica e Applicazioni,
Università di Milano "Bicocca"
Via Bicocca degli Arcimboldi 8, 20126 Milano, Italia

ABSTRACT. We present simple proofs of transience / recurrence for certain card shuffling models, that is, random walks on the infinite symmetric group.

1. CARD SHUFFLING MODELS

In this note, we consider several models of shuffling an infinite deck of cards. One of these models has been considered previously by Lawler [La]; our methods (using flows, shorting and comparison of Dirichlet forms) will – partially – simplify his result.

Card shuffling is formalized by performing successive i.i.d. random permutations in the group S_∞ of all permutations of the positive integers $\mathbb{N}$ that leave all but finitely many elements fixed. We identify the symmetric group S_n with the subgroup of S_∞ fixing all elements $> n$, so that S_∞ is the union of the S_n. We read the product of two permutations x, y from left to right, that is, $x \cdot y$ sends $j \in \mathbb{N}$ to $y\big(x(j)\big)$. Let μ be a symmetric probability measure on S_∞ whose support generates the whole group. Associate with it a sequence X_n, $n \geq 1$, of i.i.d. S_∞-valued random variables with common distribution μ, and consider the associated *random walk* $Z_n = X_1 \cdots X_n$. This means that we start with the deck of cards in order (Z_0 is the identity), and at each step we choose a random permutation X_n according to μ which tells us how to shuffle once more what we had obtained previously.

The question addressed here is the following. Will the deck of cards eventually return to its original order with probability one *(recurrence),* or is this probability strictly smaller than 1 *(transience)* ?

We now describe four shuffling models, that is, probabilities $\mu_1, \dots, \mu_4$, each one governing another random walk. We start by considering a probability distribution $p(\cdot)$ on $\{2, 3, \dots\}$.

Mathematics Subject Classification: 60J15.
Key words and phrases. Infinite permutation group, random walk, transience, recurrence.

1.) Choose first n with probability $p(n)$ and then $j \in \{1, \dots, n-1\}$ with probability $1/(n-1)$, and exchange the n-th with the j-th card. Writing $t(n,j)$ for the transposition of n and j, we have

$$\mu_1 = \sum_{n=2}^{\infty} \sum_{j=1}^{n-1} \frac{p(n)}{n-1} \, \delta_{t(n,j)} \,.$$

2.) Choose n with probability $p(n)$, and exchange the n-th with the first card. Thus

$$\mu_2 = \sum_{n=2}^{\infty} p(n) \, \delta_{t(n,1)} \,.$$

3.) Choose first n with probability $p(n)$ and then, with probability $1/2$ each, either move the n-th card to the top or the top card next below to the n-th card. Writing $c(n)$ for the cyclic permutation $(1, 2, \cdots, n)$ and $c(-n)$ for its inverse, we have

$$\mu_3 = \sum_{n=2}^{\infty} \frac{p(n)}{2} \left(\delta_{c(n)} + \delta_{c(-n)} \right) .$$

4.) Choose first n with probability $p(n)$ and then $j \in \{1, \dots, n-1\}$ with probability $1/(n-1)$, and put the cards in positions $j+1, \dots, n$ (in the same order) on top of the deck. Thus,

$$\mu_4 = \sum_{n=2}^{\infty} \sum_{j=1}^{n-1} \frac{p(n)}{n-1} \, \delta_{c(n)^j} \,.$$

Of course, δ_x indicates the Dirac mass at $x \in S_\infty$. As a consequence of what we shall prove below, we obtain the following.

Theorem. *If $p(n) \sim c/(n-\alpha)!$, then each of the above random walks is recurrent if $\alpha \le 2$ and transient if $\alpha > 2$.*

Here, $\sim$ means that quotients tend to one, c is the proper normalization constant, and for non-integer α, the factorial is defined via the Gamma function.

We remark that there is a fifth model, which is much easier but less interesting, namely, when μ_5 is equidistributed on each of $S_n \setminus S_{n-1}$, with $\mu_5(S_n \setminus S_{n-1}) = p(n)$.

Lawler [La] has considered μ_2 (he also assigns positive mass to the identity *id*, which makes no difference). For proving transience, we have found it easier to start with μ_1. We use the method of shorting, wich goes back to Nash-Williams [NW] and the flow criterion of Yamasaki [Ya] - Lyons [Ly]. We then adapt comparison techniques of Dirichlet forms, that have been successfully applied, for example, to random walks on finite groups by Diaconis and Saloff-Coste [D-S], in order to deal with the other three models. We shall need that all $p(n)$ are positive and satisfy certain monotonicity conditions which hold in the special case of the above Theorem.

2. A general recurrence criterion of Lawler

In the context of model 2, Lawler [La] has proved a recurrence criterion which adapts with very little additional effort to arbitrary symmetric random walks on direct limits of arbitrary finite groups. For convenience of the reader, we state and prove the general version.

Let $G = \bigcup_{n\geq 1} G_n$, where the G_n are finite groups and each G_n is a (proper) subgroup of G_{n+1}. On G, consider a symmetric probability measure μ whose support generates G. We may assume that $\mu(G_1) > 0$. Again, the random walk is $Z_n = X_1 \cdots X_n$, where the X_k are i.i.d. G-valued with distribution μ.

Proposition 1. *The random walk on G with law μ is recurrent, if*

$$\sum_{n=1}^{\infty} \frac{1}{|G_n|\big(1-\mu(G_n)\big)} = \infty\,.$$

Proof. Recurrence is equivalent with $\sum_n \mathbb{P}[Z_n = id] = \infty$, and $\mathbb{P}[Z_n = id] = \mu^{(n)}(id)$, where $\mu^{(n)}$ is the n-th convolution power. Let $\mu_k = \frac{1}{\mu(G_k)}\mu|_{G_k}$. By symmetry and Cauchy-Schwarz, $\mu_k^{(2n)}(id) \geq 1/|G_k|$. Hence, for arbirtrary k and n,

$$\mathbb{P}[Z_{2n} = id] \geq \mathbb{P}[Z_{2n} = id, X_i \in G_k \ \forall\ i \leq 2n] = \mu_k^{(2n)}(id)\,\mu(G_k)^{2n} \geq \frac{\mu(G_k)^{2n}}{|G_k|}\,.$$

Now let $r(k) = 1/\big(1-\mu(G_k)\big)$. Then $r(k) \leq r(k+1) \to \infty$. If $2n \leq r(k)$ then we find

$$\mathbb{P}[Z_{2n} = id] \geq \frac{1}{|G_k|}\left(1 - \frac{1}{r(k)}\right)^{r(k)} \geq \frac{c_0}{|G_k|}\,,$$

where $c_0 > 0$. Next, let $s(k) = \lfloor r(k)/2 \rfloor$. Then

$$\sum_{n=0}^{\infty} \mathbb{P}[Z_n = id] \geq \sum_{k=2}^{\infty} \sum_{\substack{n:\\ s(k-1)<n\leq s(k)}} \mathbb{P}[Z_{2n} = id] \geq c_0 \sum_{k=2}^{\infty} \frac{s(k)-s(k-1)}{|G_k|}\,. \tag{A}$$

Note that $|G_{k+1}| \geq 2|G_k|$, whence $2|G_k|^{-1} \geq \sum_{i\geq k} |G_i|^{-1}$. Therefore the right hand term in (A) is

$$\geq \frac{c_0}{2} \sum_{i=2}^{\infty} \sum_{k=2}^{i} \frac{s(k)-s(k-1)}{|G_i|} = \frac{c_0}{2} \sum_{i=2}^{\infty} \frac{s(i)-s(1)}{|G_i|} \geq \frac{c_0}{2} \sum_{i=2}^{\infty} \frac{s(i)}{|G_i|} - \frac{c_0\, s(1)}{4}\,.$$

Now, if $\sum_n r(n)/|G_n| = \infty$, then also $\sum_i s(i)/|G_i| = \infty$, and the random walk is recurrent. □

We remark that special cases of Propopsition 1 had been proved previously by Flatto and Pitt [F-P] (when each G_n is cyclic, or when G is the direct sum of finite abelian groups), and before that by Darling and Erdös [D-E] (when G is the direct sum of infintely many copies of the two-element group).

3. Un-shorting and flows

Returning to our card-shuffling models, we now study transience of μ_1.

Proposition 2. *The shuffling model no. 1 is transient, provided that*

$$\sum_{n=1}^{\infty} \frac{1}{|S_n|\, p(n+1)} < \infty .$$

We use well known methods of infinite network theory, see the survey of Woess [Wo]. Let X be an infinite, connected graph with symmetric neighbourhood relation $\sim$. With each edge $[x,y]$, we associate a *conductance* $c(x,y) = c(y,x) > 0$, such that $m(x) = \sum_{y\sim x} c(x,y) < \infty$ for all x. The associated reversible Markov chain (random walk) on X has transition probabilities $p(x,y) = c(x,y)/m(x)$. If we start with a symmetric $p(\cdot,\cdot)$, then $c(x,y) = p(x,y)$ and $m(x) = 1$. For a function $f : X \to \mathbb{R}$, its *Dirichlet sum* is

$$D(f) = \frac{1}{2} \sum_{x\sim y} c(x,y) \big(f(y) - f(x)\big)^2 .$$

The space $\mathcal{D}(X) = \{f : D(f) < \infty\}$ with $\|f\|^2 = D(f) + f(o)^2$ is Hilbert and independent of the choice of the reference vertex o. We write $\ell_0(X)$ for the finitely supported functions and $\mathcal{D}_0(X)$ for their closure in $\mathcal{D}(X)$.

Shortening is the following procedure. Let $X = \bigcup_{i\in\mathcal{I}} X_i$ be a partition such that $\mathbf{1}_{X_i} \in \mathcal{D}_0(X)$ for all i. The shorted network is the one over $\mathcal{I}$ where $i \sim j$ if $i \neq j$ and there is some edge between X_i and X_j, with conductances $c'(i,j) = \sum_{x\in X_i, y\in X_j, x\sim y} c(x,y)$. The condition on $\mathbf{1}_{X_i}$ guarantees that $m'(i) = \sum_{j\sim i} c'(i,j) < \infty$. Then one has the following recurrence criterion:

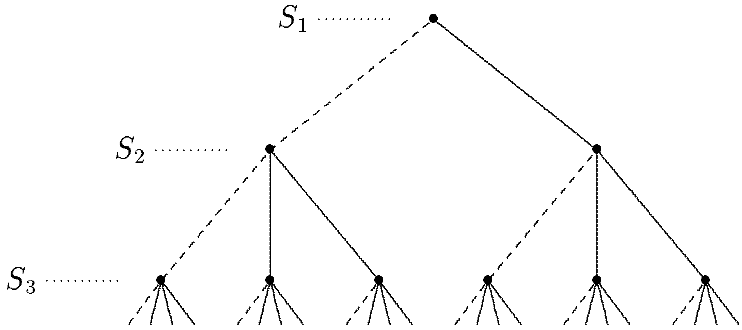
If the random walk on $\mathcal{I}$ is recurrent then so is the random walk on X. (B)

In the case when the shorted network is a half-line, this is Nash-Williams' criterion [NW]. Another well known criterion is the following:

If a network is recurrent then every sub-network is recurrent. (C)

We use these two criteria in the reversed direction. We first construct a locally finite tree with associated conductances given by the $p(n)$ and show that it is transient. We then short it to obtain a sub-network of the one associated with μ_1 on S_∞.

The vertex set of our tree T is the *disjoint* union of the S_n. The root is the identity, and edges occur only between S_{n-1} and S_n, namely from $x \in S_{n-1}$ to $x \cdot t(n,j)$, where $j < n$, and from $x \in S_{n-1}$ to the copy of $x = x \cdot id$ in S_n. We call the latter the *improper* edges; they are dashed in the figure below.



T is indeed a tree because of the fact that for each $y \in S_n \setminus S_{n-1}$ there are uniqe $j < n$ and $x \in S_{n-1}$ such that $y = x \cdot t(n,j)$. To each edge from S_{n-1} to S_n in T (proper or improper) we assign conductance $p(n)/(n-1)$.

Proof of Proposition 1. We use flows. On a network $\big(X, c(\cdot,\cdot)\big)$, a *unit flow* from o to ∞ is a function $F : X \times X$ such that $F(x,y) \neq 0$ only when x, y are neighbours, $F(x,y) = F(y,x)$, and $\sum_{y:y\sim x} F(x,y) = 0$ for all $x \neq o$, while $\sum_{y:y\sim o} F(o,y) = 1$. Its *energy* is $\sum_{x\sim y} F(x,y)^2/c(x,y)$. The *flow criterion* of Yamasaki [Ya] and Lyons [Ly] is the following.

The random walk on X is transient if and only if there is a (unit) flow from o to ∞ with finite energy.

We now use the unit flow from $o = id \in S_1$ that branches equally at each vertex when moving down from o. That is, $F(x,y) = 1/n!$ when $x \in S_{n-1}$ and $y \in S_n$ are neighbours in T (properly or improperly). It is an immediate exercise to verify that F has finite energy precisely when the condition of Proposition 2 holds, thus proving transience of the random walk on T.

If $x \in S_\infty$ then $x \in S_k$ for a minimal $k = k(x)$, and in T, there is a ray (half-line) R_x of successive improper edges starting at x (the latter viewed as a vertex in $S_{k(x)}$). We now short each of these *improper rays* to a single point, again denoted x. Let $R_x(n)$ be the finite path obtained by truncating R_x at level $n \geq k(x)$. Then $D(\mathbf{1}_{R_x} - \mathbf{1}_{R_x(n)}) = \frac{p(n)}{n-1} + \sum_{k>n} p(k) \to 0$, whence the indicator function of each improper ray is in $\mathcal{D}_0(T)$, and shorting is legitimate. By (B), the random walk on the shorted network is also transient. This is in fact a spanning tree of the Cayley graph of S_∞ with respect to the generating set consisting of all transpositions, and it is certainly a subnetwork of the network corresponding to the random walk with law μ_1. Our tree T arises from *un-shorting* this subnetwork. By (C), the latter must also be transient. □

We remark that writing down this proof takes more time than explaining it with the figure at hand. In particular, the three ingredients (flows, shorting, passing to sub-networks) have by now become part of the random walk folklore. The critical eye might observe that we could have constructed the finite energy flow directly on the shorted tree. We think that the above approach is nicer. Also, it seems that "un-shorting" has not been used before for proving transience.

4. Comparison of Dirichlet forms

In order to deduce from Proposition 1 transience criteria for models 2, 3 and 4, we use the following method. Let X be equipped with two different graph structures, associated conductances and corresponding transition matrices P and Q. We write D_P and D_Q for the respective Dirichlet sums.

If $D_Q(f) \leq c\, D_P(f)$ for all $f \in \ell_0(X)$, where $c > 0$, then transience of Q implies transience of P. (D)

Let us adapt this to the context of a group G. If $a \in G$ then we write

$$D_a(f) = \sum_{x\in G} \big(f(x\cdot a) - f(x)\big)^2 ,$$

and if μ is a symmetric probability measure on G, then $D_\mu(f) = \sum_{a\in G}\mu(a)\,D_a(f)$. The graph structure associated with μ is the Cayley graph of G with respect to $S = \mathrm{supp}(\mu)$. Note that D_μ is is not precisely the corresponding Dirichlet sum as defined in §2, because edges $[x, x\cdot a]$ are counted twice when $a \in S$ is not idempotent. This does not matter when applying (D), as the factor $1/2$ which is eventually missing can be subsumed in the constant c. Basic tool is the following.

Lemma. *If $a = b_1 \cdots b_k$ then $D_a(f) \le k\Big(D_{b_1}(f) + \cdots + D_{b_k}(f)\Big)$.*

Proof. This is straightforward via the Cauchy-Schwarz inequality and the expansion $f(x\cdot a) - f(x) = \sum_{i=1}^k \big(f(x\cdot y_i) - f(x\cdot y_{i-1})\big)$, where $y_i = b_1\cdots b_i$. □

Proposition 3. *If $p(n) > 0$ for all $n \ge 2$ and $\limsup p(n+1)/p(n) < 1$ then there are $c_2, c_4 > 0$ such that $D_{\mu_1} \le c_2\, D_{\mu_2}$ and $D_{\mu_1} \le c_4\, D_{\mu_4}$.*

Proof. Model 2. If $2 \le j \le n-1$ then $t(n,j) = t(j,1)\cdot t(n,1)\cdot t(j,1)$, and the Lemma yields $D_{t(n,j)} \le 6\,D_{t(j,1)} + 3\,D_{t(n,1)}$. Therefore

$$
\begin{aligned}
D_{\mu_1} &\le \sum_{n=2}^{\infty}\frac{p(n)}{n-1}D_{t(n,1)} + \sum_{n=3}^{\infty}\sum_{j=2}^{n-1}\left(6\frac{p(n)}{n-1}\,D_{t(j,1)} + 3\frac{p(n)}{n-1}\,D_{t(n,1)}\right)\\
&\le 6\sum_{n=2}^{\infty}\left(p(n) + \sum_{k=n+1}^{\infty}\frac{p(k)}{k-1}\right)D_{t(n,1)}\,.
\end{aligned}
$$

The assumptions on $p(n)$ yield that $\frac{p(n+k)}{n+k-1} \le C\,\lambda^k\,\frac{p(n)}{n-1}$ for all n,k, where $C \ge 1$ and $0 < \lambda < 1$. Therefore

$$
\sum_{k=n+1}^{\infty}\frac{p(k)}{k-1} \le \bar C\,\frac{p(n)}{n-1} \tag{E}
$$

with $\bar C = C\,\lambda/(1-\lambda)$, and we find $c_2 = 1 + \bar C$.

Model 4. We have $t(n,j) = c(-n)^j\cdot c(n-1)^j\cdot c(-j)$ when $n > j \ge 2$, while $t(n,1) = c(-n)\cdot c(n-1)$ for $n \ge 3$, and $t(2,1) = c(2)$. Observe that $D_{a^{-1}}(f) = D_a(f)$. Therefore

$$
\begin{aligned}
D_{\mu_1} &\le p(2)\,D_{c(2)} + 2\sum_{n=3}^{\infty}\frac{p(n)}{n-1}\Big(D_{c(n)} + D_{c(n-1)}\Big)\\
&\qquad + 3\sum_{n=3}^{\infty}\sum_{j=2}^{n-1}\frac{p(n)}{n-1}\Big(D_{c(n)^j} + D_{c(n-1)^j} + D_{c(j)}\Big)\\
&= \left(p(2) + p(3) + 3\sum_{k=3}^{\infty}\frac{p(k)}{k-1}\right)D_{c(2)}\\
&\qquad + \sum_{n=3}^{\infty}\left(2\frac{p(n)}{n-1} + 2\frac{p(n+1)}{n} + 3\sum_{k=n+1}^{\infty}\frac{p(k)}{k-1}\right)D_{c(n)}\\
&\qquad + \sum_{n=3}^{\infty}\sum_{j=2}^{n-1}\left(3\frac{p(n)}{n-1} + 3\frac{p(n+1)}{n}\right)D_{c(n)^j}\,.
\end{aligned}
$$

Once more using (E), we find $c_4 = 3 + 3C/(1-\lambda)$. □

Proposition 4. *If $p(2) > 0$ then there is $c_3 > 0$ such that* $D_{\mu_2} \le c_3\, D_{\mu_3}$.

Proof. We have $t(2,1) = c(2)$. If $n \ge 3$ then $t(n,1) = c(n) \cdot c(2) \cdot c(-n)$. Therefore

$$\begin{aligned} D_{\mu_2} &\le p(2)\, D_{c(2)} + \sum_{n=3}^{\infty} \Big(6p(n)\, D_{c(n)} + 3p(n)\, D_{c(2)} \Big) \\ &= \Big(3 - 2p(2)\Big)\, D_{c(2)} + 6 \sum_{n=3}^{\infty} p(n)\, D_{c(n)}\,, \end{aligned}$$

which is bounded above by $\max\{6\,,\ 3/p(2)\}\, D_{\mu_3}$. □

Combining Propositions 2–4 with criterion (D), we now obtain the following.

Corollary. *If* $\sum_{n=1}^{\infty} 1/\big(|S_n|\, p(n+1)\big) < \infty$ *and* $\limsup p(n+1)/p(n) < 1$ *then the shuffling models no. 2 – 4 are transient.*

In particular, when $\limsup p(n{+}1)/p(n) < 1$ then the condition of Proposition 2 is necessary and sufficient for transience. The Theorem stated in the introduction arises as a special case.

In Proposition 3 and the Corollary, one can of course replace the condition $\limsup p(n+1)/p(n) < 1$ by some weaker hypothesis that imples (E), or by (E) itself.

Lawler [La] has proved that for transience of model no. 2, already the first of the two conditions in the above Corollary is sufficient. In view of Proposition 4, the lim sup-condition can also be dropped when considering transience of model no. 3.

We conclude with an outlook. If one wants to reduce the gap between the criteria of Propositions 1 and 2, one possible approach will be to use the representation theory of S_∞, or the S_n, respectively. Denote by σ_n the uniform distribution on $\{t(n,j) : j < n\}$. Then the σ_n commute under convolution, which is why model 1 is preferable to the others in this context, compare with Diaconis [Di]. We intend to come back to this in future work.

References

[D-E] D. A. Darling and P. Erdös, *On the recurrence of a certain chain*, Proc. Amer. Math. Soc. **19** (1968), 336–338.

[Di] P. Diaconis, *Group representations in Probability and Statistics*, IMS, Hayward, California, 1988.

[D-S] P. Diaconis and L. Saloff-Coste, *Comparison techniques for random walk on finite groups*, Ann. Probab. **21** (1993), 2131–2156.

[F-P] L. Flatto and J. Pitt, *Recurrence criteria for random walks on countable abelian groups*, Illinois J. Math. **18** (1974), 1–19.

[La] G. F. Lawler, *Recurrence and transience for a card shuffling model*, Comb. Probab. Comp. **4** (1995), 133–142.

[Ly] T. Lyons, *A simple criterion for transience of a reversible Markov chain*, Ann. Probab. **11** (1983), 393–402.

[NW] C. St. J. A. Nash-Williams, *Random walks and electric currents in networks*, Proc. Cambridge Phil. Soc. **55** (1959), 181–194.

[Wo] W. Woess, *Random walks on infinite graphs and groups: a survey on selected topics*, Bull. London Math. Soc. **26** (1994), 1–60.

[Ya] M. Yamasaki, *Discrete potentials on an infinite network*, Mem. Fac. Sci. Shimane Univ. **13** (1979), 31–44.