A Better Way to Deal the Cards

Mark A. Conger and Jason Howald

Abstract. Most of the work on card shuffling assumes that all the cards in a deck are distinct, and that in a well-shuffled deck all orderings need to be equally likely. We consider the case of decks with repeated cards and decks which are dealt into hands, as in bridge and poker. We derive asymptotic formulas for the randomness of the resulting games. Results include the influence of where a poker deck is cut, and the fact that switching from cyclic dealing to back-and-forth dealing will improve the randomness of a bridge deck by a factor of 13.

1. INTRODUCTION. Card shuffling has been used as an example of a mixing problem since the early part of the twentieth century. The fundamental question may be boiled down to: "How fast does repeated shuffling randomize the deck?" We will explore how to make mathematical sense of that question below.

At the same time, in many card games, such as bridge, euchre, and straight poker, the players receive "hands" of cards which are dealt from the deck after it has been shuffled. A hand has no inherent order; in other words, the sequence in which the cards arrive in front of a player is unimportant. Thus the process of shuffling and dealing partitions the deck into sets of cards of predetermined sizes, and our goal should be to make the partition as unpredictable as possible. This might not require fully randomizing the deck.

For instance, bridge is a game played with 4 players and a 52-card deck. Each player receives 13 cards. The usual method is to shuffle and then deal cyclically: first card to the player on the dealer's left, next card to the player on *his* left, and so on, clockwise around the table. Why deal this way, instead of just giving 13 cards off the top of the shuffled deck to the first player, then 13 to the next player, and so on? If the deck were perfectly randomized by the shuffling, the method of dealing would not matter.

One reason not to simply "cut the deck into hands" is to prevent a dishonest dealer from stacking the deck through unfair shuffling. But one might also guess, correctly, that dealing cyclically augments the randomness of the game when the dealer is honest but has not shuffled the deck thoroughly. In this paper we address the questions: "How much difference does a dealing method make?" and "Is cyclic dealing the best method?" The goal is to show that the answers are "Quite a bit" and "No, we can do a lot better!" We present results for bridge and straight poker as examples. We also show how to estimate the randomness in a deck that has distinct cards and a deck with only two types of cards.

2. A BRIEF HISTORY OF CARD SHUFFLING. Henri Poincaré devoted eight sections of his 1912 book *Calcul des Probabilités* [**16**, §225–232] to card shuffling. He did not attempt to model any particular kind of shuffling, but showed that any shuffling method which meets certain mild criteria, if applied repeatedly to a deck, will eventually result in a well-mixed deck—that is, with enough shuffles the bias can be made arbitrarily small. About the same time Markov [**15**] was creating the more general theory of Markov chains, and he often used card shuffling as an example. The verdict of history seems to be that Markov justly deserves credit for the theory named

doi:10.4169/000298910X515758

after him, but that Poincaré anticipated some of Markov's ideas in his work on card shuffling. Most subsequent work on shuffling has approached the problem as a Markov chain.

In the 1950s Gilbert and Shannon [**12**] considered the problem of **riffle shuffling**. Riffle shuffling is the most common method used by card players to randomize a deck: the shuffler cuts the deck into two packets, then interleaves (riffles) them together in some fashion. Using the new science of information theory, Gilbert and Shannon began the inquiry into how fast riffle shuffling mixes a deck. In the 1980s Reeds [**17**] and Aldous [**1**] added the assumption that all cut/riffle combinations are equally likely, and that has become known as the Gilbert-Shannon-Reeds or GSR model of card shuffling.

In 1992, in the most celebrated paper on card shuffling to date [**3**], Bayer and Diaconis generalized the GSR shuffle to the ***a*-shuffle**. Let a be a positive integer, and cut a deck into a packets (one imagines an a-handed dealer in a futuristic casino), then riffle them together in some fashion. Assuming as before that all cut/riffle combinations are equally likely, performing a randomly selected a-shuffle followed by a randomly selected b-shuffle turns out to be equivalent to performing a randomly selected ab-shuffle. In particular that means that a sequence of k GSR shuffles is equivalent to a single 2^k-shuffle. Thus if we understand a-shuffles we implicitly understand repeated shuffles.

Bayer and Diaconis found a remarkably simple and elegant formula for the probability of a particular permutation π after an a-shuffle, namely

$$\mathbb{P}_a(\pi) = \frac{1}{a^n}\binom{a+n-\operatorname{des}(\pi)-1}{n}, \tag{1}$$

where n is the size of the deck and

$$\operatorname{des}(\pi) := \#\{i : \pi(i) > \pi(i+1)\}$$

is the number of descents in the permutation π.

Note. There are two ways to view a permutation:

1. as a bijection π from $\{1, 2, \ldots, n\}$ to iteslf, so that π, when applied to a sequence of objects, moves the object in position i to position $\pi(i)$, and
2. as an ordering $\sigma = \sigma_1, \sigma_2, \ldots, \sigma_n$ of $1, 2, \ldots, n$.

In this paper our decks will contain repeated cards, and our permutations will act to rearrange them. So we will consistently interpret permutations as maps. This may disorient readers who are used to the other viewpoint.

In order to analyze the progress of a shuffler toward a well-mixed deck, we need a measure of how close the distribution after an a-shuffle is to the uniform distribution (all orderings equally likely). For this Bayer and Diaconis use **variation distance from uniform**, which may be defined as

$$||\mathbb{P}_a - U|| := \frac{1}{2}\sum_{\pi\in S_n} |\mathbb{P}_a(\pi) - U(\pi)|\,, \tag{2}$$

where U represents the uniform distribution on permutations, i.e., $U(\pi) = 1/n!$ for all $\pi \in S_n$. In terms of cards, the most biased game one could play with the shuffled deck is the two-player game in which player 1 wins whenever the ordering of the deck has probability higher than it should be under the uniform distribution. So $||\mathbb{P}_a - U||$

is the maximum bias toward any player in any game one might care to play with the shuffled deck. That is, the probability of any set of permutations is within $||\mathbb{P}_a - U||$ of what it would be under the uniform distribution.

Using the probability formula in (1) and the knowledge that the number of permutations in S_n with d descents is the well-studied Eulerian number $\left\langle {n \atop d} \right\rangle$ (see for example [**13**], [**4**], [**18**]), Bayer and Diaconis are able to compute

$$||\mathbb{P}_a - U|| = \frac{1}{2}\sum_{d=0}^{n-1}\left\langle {n \atop d} \right\rangle \left| \frac{1}{a^n}\binom{a+n-d-1}{n} - \frac{1}{n!}\right|$$

very quickly. The result, for $n = 52$ cards and a between 1 and 1024, is graphed in Figure 1. The horizontal scale is logarithmic to represent the fact that a 2^k-shuffle is the same as k GSR shuffles.

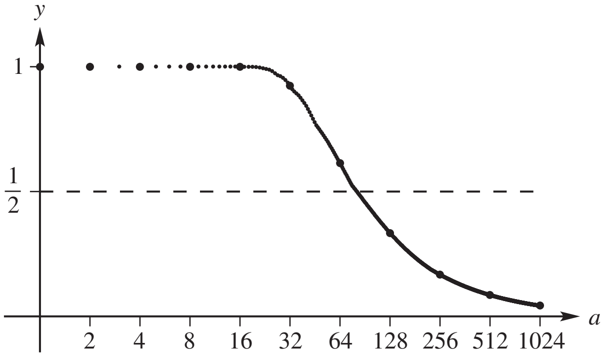


Figure 1. The variation distance from uniform of a distinct 52-card deck after an a-shuffle.

The "waterfall" shape of the graph is typical of what Aldous [**1**] calls "rapidly mixing Markov chains": negligible change after the first few GSR shuffles, followed by a fast approach to uniform ("the cutoff"), and eventually halving with each extra shuffle. Note that if we cast the problem in terms of a-shuffles, the ultimate exponential decay means the variation distance approaches κ_1/a for some constant κ_1 as a gets large.

Many more papers on card shuffling and its applications have been published. See [**10**] for a survey and [**14**] for an excellent exposition of the Bayer and Diaconis results. There are other choices besides variation distance for measuring randomness—see [**9**, pp. 22–23], [**3**, §5.1], [**5**, §5], [**19**, §3], and [**2**, §1] for alternatives.

3. REPEATED CARDS AND DEALING METHODS.

The implicit assumption in Bayer and Diaconis and most other work on card shuffling is that all cards in the deck are distinct. Here we will take up the alternative, explored recently by Conger and Viswanath [**7, 8**] and Assaf, Diaconis, and Soundararajan [**2**]. Suppose a deck D is a sequence of cards, each of which has a value taken from some fixed set of values, and we allow two cards to have the same value. This complicates the problem in the following ways:

- Decks (ordered sequences of cards) and transformations between decks can no longer be identified with permutations. Instead for each pair of decks there is a set of permutations which transform the first into the second (and a different set which goes the other way). The transformation sets are easy to describe, but it is difficult to find their probabilities after shuffling.

- The initial order of a deck, and not just its composition, affects how fast the distribution approaches uniform.

If D' is some rearrangement of D, let $T(D, D')$ be the set of permutations which, when applied to D, result in D'. ($T(D, D')$ is a left coset of the stabilizer $\mathrm{stab}(D) = T(D, D)$ and a right coset of $\mathrm{stab}(D') = T(D', D')$.) Thus the probability of obtaining D' as a result of a-shuffling D is

$$\mathbb{P}_a(D \to D') := \sum_{\pi \in T(D,D')} \mathbb{P}_a(\pi) = \frac{1}{a^n} \sum_d b_d \binom{a+n-d-1}{n},$$

where b_d is the number of permutations in $T(D, D')$ with d descents. In this context we call D the **source deck**, D' the **target deck**, and

$$\mathcal{D}(D, D'; x) := \sum_{\pi \in T(D,D')} x^{\mathrm{des}(\pi)} = \sum_d b_d x^d \tag{3}$$

the **descent polynomial** of D and D'. (The reader will kindly forgive the many different uses of the letter "D" in this paper. D will always be a source deck, D' a target deck, d an integer which represents a number of descents, and $\mathcal{D}$ the descent polynomial.)

If we give D an a-shuffle, then the variation distance between the resulting distribution on decks and the uniform distribution is

$$||\mathbb{P}_a - U|| = \frac{1}{2} \sum_{D' \in \mathcal{O}(D)} \left| \mathbb{P}_a(D \to D') - \frac{1}{N} \right|, \tag{4}$$

where $\mathcal{O}(D)$ is the set of reorderings of D (i.e., D's orbit when acted on by S_n) and N is the size of $\mathcal{O}(D)$. We will refer to this as the **fixed source** case.

On the other hand, suppose we are playing bridge. Here all the cards have distinct values; let $e(1), e(2), \ldots, e(52)$ be the initial order of the cards. The dealer a-shuffles the deck and then deals it out to the four players, who are referred to as North, East, South, and West. Since each player receives 13 cards, we can describe a method of dealing cards as a sequence of 13 `N`'s, 13 `E`'s, 13 `S`'s, and 13 `W`'s. For instance,

$$D' := \texttt{NESWNESW} \cdots \texttt{NESW} = (\texttt{NESW})^{13}$$

represents cyclic dealing, where the top card goes to North, the second to East, etc. We can describe a particular partition of the deck into hands by a string of the same type. For example,

$$D := \texttt{NSEENNWEWSSWESWNNNEESSSSSESWWNNSENWSEWSWWWEENEWNNNWE} \tag{5}$$

refers to the partition in which the North player gets cards $e(1), e(5), e(6), \ldots, e(50)$, the East player gets cards $e(3), e(4), e(8), \ldots, e(52)$, and so on. So D, like D', is a function from $\{1, 2, \ldots, 52\}$ to $\{\texttt{N}, \texttt{E}, \texttt{S}, \texttt{W}\}$, and $D(i)$ is the player who receives card $e(i)$.

In order for the North player to receive the cards assigned by D, the shuffle must move those cards to the positions occupied by `N`'s in D', and likewise for the other players. In other words, if we think of D and D' as *decks* with cards of value `N`, `E`, `S`, and `W`, the shuffle will produce the desired partition if and only if it takes D to D'.

Thus we already have a notation for the distribution over partitions that an a-shuffle followed by the dealing method D' produces, and the variation distance between that distribution and uniform is

$$||\mathbb{P}_a - U|| = \frac{1}{2} \sum_{D \in \mathcal{O}(D')} \left| \mathbb{P}_a(D \to D') - \frac{1}{N} \right|, \tag{6}$$

where N is now the number of possible partitions. This is the **fixed target** case. Note that despite the strong similarity between (4) and (6), fixed source and fixed target are dual problems, not identical, because the transition probability $\mathbb{P}_a(D \to D')$ is not symmetric.

4. CALCULATING TRANSITION PROBABILITIES. Unfortunately, computing $||\mathbb{P}_a - U||$ precisely for realistically sized decks is prohibitively complicated in both the fixed source and fixed target cases. Both cases require knowledge of the transition probabilities $\mathbb{P}_a(D \to D')$ in order to calculate variation distance from uniform. Conger and Viswanath [**7**] showed that calculation of transition probabilities is computationally equivalent to calculation of the coefficients of the descent polynomial $\mathcal{D}(D, D'; x)$. Unfortunately, the same authors subsequently found [**8**] that for certain decks the calculation belongs to a class of counting problems called #P, and is in fact #P-complete. As with NP-complete problems, it is generally believed that #P-complete problems do not admit efficient solutions.

Barring a method for calculating variation distance without first computing transition probabilities, the question shifts to approximation. Theorem 1 (below) will allow us to approximate transition probabilities when a is large, given some simple information about the decks. To that end, here are some definitions:

If u and v are card values, we say that D has a ***u-v* digraph at *i*** if $D(i) = u$ and $D(i+1) = v$. We say that D has a ***u-v* pair at (*i*, *j*)** if $i < j$, $D(i) = u$, and $D(j) = v$. The distinction between digraphs and pairs is akin to that between descents and inversions in a permutation. Let

$$W(D, u, v) := \#\{u\text{-}v \text{ digraphs in } D\} - \#\{v\text{-}u \text{ digraphs in } D\}, \tag{7}$$

$$Z(D, u, v) := \#\{u\text{-}v \text{ pairs in } D\} - \#\{v\text{-}u \text{ pairs in } D\}. \tag{8}$$

For example, the deck $D = \texttt{ABAAABABB}$ has 3 $\texttt{A-B}$ digraphs and 2 $\texttt{B-A}$ digraphs, so $W(D, \texttt{A}, \texttt{B}) = 3 - 2 = 1$. D has 15 $\texttt{A-B}$ pairs and 5 $\texttt{B-A}$ pairs, so $Z(D, \texttt{A}, \texttt{B}) = 15 - 5 = 10$. Note that both W and Z are antisymmetric in u and v: $W(D, u, v) = -W(D, v, u)$ and $Z(D, u, v) = -Z(D, v, u)$.

In the theorem below we assume that D is a deck of n cards. For each card value v, n_v is the number of cards in D with that value. For convenience we assume an implicit order on the values; the particular order chosen is arbitrary and does not affect the result.

Theorem 1. *Suppose D is as above, and D' is a reordering of D. Then*

$$\mathbb{P}_a(D \to D') = \frac{1}{N} + c_1(D, D')a^{-1} + \mathrm{O}(a^{-2}),$$

where N is the number of reorderings of D and

$$c_1(D, D') = \frac{n}{2N} \sum_{u<v} \frac{W(D, u, v) Z(D', u, v)}{n_u n_v}. \tag{9}$$

We begin with a plausibility argument in favor of formula (9). An a-shuffle begins with an a-cut, which arranges the cards into a piles. If a is very large, most piles will have size 0 or 1. If no pile has two or more cards, then no ordering information will survive the shuffle. The next most likely case is that the a-cut produces just one pile with two cards u and v. Should this happen, these cards must come from a u-v digraph in D. When the cards are reassembled, these two must remain in order, becoming a u-v pair in D'. Thus the main source of bias in $\mathbb{P}_a(D \to D')$ is the relationship between digraphs in D and pairs in D'. This suggests the formula $W(D, u, v)Z(D', u, v)$.

Proof of Theorem 1. Let S be the size of the stabilizer of D in S_n. Since $T(D, D')$ is a coset of the stabilizer, its size is S also, and since there are N such cosets, $NS = n!$. Let b_d be the number of permutations in $T(D, D')$ with d descents; so $\sum_d b_d = S$. Then

$$\begin{aligned}
\mathbb{P}_a(D \to D') &= \frac{1}{a^n} \sum_d b_d \binom{n+a-d-1}{n} \\
&= \frac{1}{a^n} \sum_d b_d \frac{(a-d)(a-d+1)\cdots(a-d+n-1)}{n!} \\
&= \frac{1}{n!} \sum_d b_d \left(1 + \frac{-d}{a}\right)\left(1 + \frac{1-d}{a}\right)\cdots\left(1 + \frac{n-1-d}{a}\right) \\
&= \frac{1}{NS}\left(\sum_d b_d + a^{-1} \sum_d b_d \left(\frac{n(n-1)}{2} - nd\right) + \mathrm{O}(a^{-2})\right).
\end{aligned}$$

So the constant term is $1/N$, which is what we expect: it means that if we shuffle D for long enough, the probability of obtaining any particular deck approaches $1/N$, i.e., the distribution on decks approaches uniform. The coefficient of a^{-1} is

$$\frac{n}{2NS} \sum_d b_d(n-1-2d) = \frac{n}{2N}\mathbb{E}(n-1-2\operatorname{des}(\pi)),$$

where π is a permutation chosen uniformly from $T(D, D')$ and $\mathbb{E}$ represents expectation. Recall that descents are positions i with $1 \leq i \leq n-1$ such that $\pi(i) > \pi(i+1)$. The other positions are **ascents**, and if we denote their number by $\operatorname{asc}(\pi)$ then we must have $\operatorname{des}(\pi) + \operatorname{asc}(\pi) = n-1$. So the number we are after is

$$\frac{n}{2N}\mathbb{E}(\operatorname{asc}(\pi) - \operatorname{des}(\pi)) = \frac{n}{2N} \sum_{i=1}^{n-1} \mathbb{E}\omega_i(\pi), \tag{10}$$

where

$$\omega_i(\pi) = \begin{cases} 1 & \text{if } \pi(i) < \pi(i+1), \\ -1 & \text{if } \pi(i) > \pi(i+1). \end{cases}$$

Let the first card in D have value u and the second have value v. Suppose first that $u = v$. π must take those two cards to two positions in D' which have value u, but otherwise it has no reason to prefer any particular destinations; thus it is equally likely that $\pi(1) < \pi(2)$ as that $\pi(1) > \pi(2)$. So if $u = v$ then $\mathbb{E}\omega_1(\pi) = 0$.

On the other hand, suppose $u \neq v$. Then π picks a destination for the top card uniformly from among those j for which $D'(j) = u$, and likewise $\pi(2)$ is chosen uniformly and independently from $D'^{-1}(v)$. We will have $\omega_1(\pi) = 1$ if the first choice is less than the second, that is, if π maps $\{1, 2\}$ to a u-v pair in D'. $\omega_1(\pi)$ will be -1 if π maps $\{1, 2\}$ to a v-u pair in D'. Each pair is equally likely, so

$$\mathbb{E}\omega_1(\pi) = \frac{\#\{u\text{-}v \text{ pairs in } D'\} - \#\{v\text{-}u \text{ pairs in } D'\}}{\#\{u\text{-}v \text{ pairs in } D'\} + \#\{v\text{-}u \text{ pairs in } D'\}} = \frac{Z(D', u, v)}{n_u n_v}.$$

All the other ω_i are calculated in the same way, and if we group them according to the values of the digraphs in D, we get the desired result for c_1. ∎

5. THE FIXED SOURCE CASE. Now that we can approximate the transition probability between two decks, we can approximate variation distances. In the case of a fixed source deck D we have

$$||\mathbb{P}_a - U|| = \frac{1}{2} \sum_{D' \in \mathcal{O}(D)} \left| \mathbb{P}_a(D \to D') - \frac{1}{N} \right| = \kappa_1 a^{-1} + \mathrm{O}(a^{-2}),$$

where

$$\kappa_1 = \kappa_1(D) := \frac{1}{2} \sum_{D' \in \mathcal{O}(D)} \left| c_1(D, D') \right|. \tag{11}$$

So $\kappa_1(D)$ is a measure of how hard it is to shuffle the deck D.

5.1. All-Distinct Decks. For example, consider shuffling a deck D containing n distinct cards. Without loss of generality we may assume that $D = 1, 2, \ldots, n$, i.e., that the value of the card in position i is i. Each reordering of D is produced by a unique permutation, so

$$\kappa_1(D) = \frac{1}{2} \sum_{\pi \in S_n} |c_1(D, \pi D)|.$$

Since each card appears once and all the digraphs in D are of the form $(i, i+1)$ we can reduce (9) to

$$c_1(D, \pi D) = \frac{n}{2N} \sum_{i=1}^{n-1} Z(\pi D, i, i+1).$$

$Z(\pi D, i, i+1)$ is 1 if π has an ascent at i and -1 if π has a descent at i, so we have

$$\kappa_1(D) = \frac{n}{4N} \sum_{\pi \in S_n} |\mathrm{asc}(\pi) - \mathrm{des}(\pi)| = \frac{n}{4N} \sum_d \left\langle {n \atop d} \right\rangle |n - 1 - 2d|. \tag{12}$$

The Eulerian numbers can be calculated using a simple recurrence [**13**], so (12) is all we need to find the long-term behavior of variation distance from uniform after shuffling a deck of n distinct cards. For instance, when $n = 52$ we find that κ_1 is

$$\frac{146020943891326775423340146124729913263177343486982212261189487693}{331435631044312453039368165912244275868217888892518400000000000000},$$

which is approximately 44.06. (In general, κ_1 is about $n\sqrt{(n+1)/24\pi}$ for a deck of n distinct cards [**6**].) So after giving a deck of 52 distinct cards an a-shuffle, where a is large, the variation distance from uniform will be approximately $44.06/a$. We can compare that with the exact results Bayer and Diaconis obtained for the same deck, to see how big a has to be to make the approximation a good one. Figure 2 shows the result.

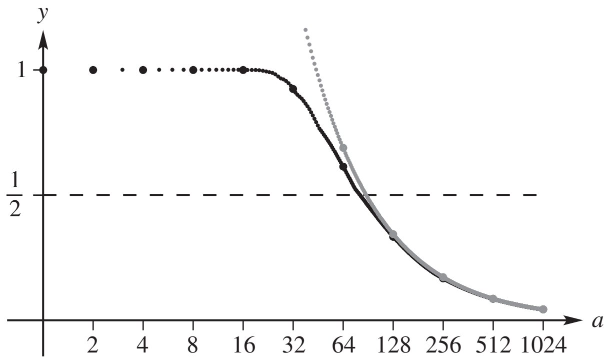
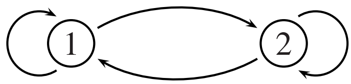


Figure 2. The variation distance from uniform, and first-order approximation, of a distinct 52-card deck after an a-shuffle. The actual variation distance is graphed in black, and the first order approximation κ_1/a, with $\kappa_1 = 44.05710497$, is graphed in gray.

5.2. Decks with Two Card Types. Let G_k be a complete directed graph on k vertices, with loops at each vertex. A deck with cards of k different types may be thought of as a walk on G_k, where the starting position is the top card and the ending position is the bottom card. Each edge in the walk represents a digraph in the deck.



Consider a deck D with only two types of cards, which we will unimaginatively label 1 and 2. If the deck begins with 1 and ends with 2, then the corresponding walk must have traversed the 1-2 edge once more than the 2-1 edge, and therefore $W(D, 1, 2) = 1$. Likewise beginning with a 2 and ending with a 1 makes $W(D, 1, 2) = -1$. But beginning and ending with the same type of card means that both edges were traversed the same number of times, so in that case $W(D, 1, 2)$ is 0. With only two types of cards (9) reduces to

$$c_1(D, D') = \frac{n}{2n_1 n_2 N} W(D, 1, 2) Z(D', 1, 2),$$

which vanishes if $W(D, 1, 2)$ is 0, for all D'. So if the top and bottom cards of the unshuffled deck are the same, κ_1 will be 0, meaning that in the long run the variation distance decreases at least as fast as some multiple of a^{-2}. So we have the surprising result that a shuffler of this type of deck can greatly speed the deck's approach to randomness by making sure that the top and bottom cards are the same before he begins shuffling.

We will concentrate on the fixed target case for the rest of this paper, but the interested reader may consult [**6**] for more results in the fixed source case.

6. THE FIXED TARGET CASE (DEALING INTO HANDS). If we fix the target deck (i.e., dealing method) D', we have

$$||\mathbb{P}_a - U|| = \frac{1}{2} \sum_{D \in \mathcal{O}(D')} \left| \mathbb{P}_a(D \to D') - \frac{1}{N} \right| = \overline{\kappa}_1 a^{-1} + \mathrm{O}(a^{-2}),$$

where

$$\overline{\kappa}_1 = \overline{\kappa}_1(D') := \frac{1}{2} \sum_{D \in \mathcal{O}(D')} \left| c_1(D, D') \right|.$$

So $\overline{\kappa}_1(D')$ is a measure of how much randomness a dealing method contributes. The smaller $\overline{\kappa}_1$ is, the fairer the deal. The sum above seems intractably large, so it is useful to have an alternative algorithm for the calculation of $\overline{\kappa}_1(D')$. First notice that

$$\begin{aligned}
\overline{\kappa}_1(D') &= \frac{1}{2} \sum_{D \in \mathcal{O}(D')} \left| c_1(D, D') \right| \\
&= \frac{1}{2} \sum_{D \in \mathcal{O}(D')} \left| \frac{n}{2N} \sum_{u<v} \frac{W(D, u, v) Z(D', u, v)}{n_u n_v} \right| \\
&= \frac{n}{4N} \sum_{D \in \mathcal{O}(D')} \left| \sum_i \frac{Z(D', D(i), D(i+1))}{n_{D(i)} n_{D(i+1)}} \right| \\
&= \frac{n}{4N} \sum_{D \in \mathcal{O}(D')} |\theta(D)|,
\end{aligned}$$

where

$$\theta(D) := \sum_i \frac{Z(D', D(i), D(i+1))}{n_{D(i)} n_{D(i+1)}}.$$

For our cases, there are far fewer possible values for $\theta(D)$ than there are decks $D \in \mathcal{O}(D')$, so we want to reason about the distribution of values of $\theta(D)$ as D ranges over $\mathcal{O}(D')$. To prepare a recursion, we will need this distribution to depend also on the last card of D, and we will need to consider decks with fewer cards than D. Since $\mathcal{O}(D')$ depends on the number of cards of each type in D', but not on their order, let $\overline{m} = (m_1, m_2, \ldots, m_k)$ be an integer vector, representing a collection of $m_1 \le n_1$ cards labeled 1, $m_2 \le n_2$ cards labeled 2, etc. If $v \in \{1, \ldots, k\}$ is a card value, we write $D \dashv v$ to mean that the last card of D is v, and $D \dashv uv$ to mean that D ends with the digraph uv. We also write:

$$\begin{aligned}
\mathcal{O}(\overline{m}) &:= \begin{cases} \mathcal{O}(1^{m_1} 2^{m_2} \cdots k^{m_k}) & \text{if all } m_i \ge 0, \\ \emptyset & \text{otherwise,} \end{cases} \\
g_{\overline{m},v}(t) &:= \sum_{\substack{D \in \mathcal{O}(\overline{m}) \\ D \dashv v}} t^{\theta(D)}.
\end{aligned}$$

Then $\sum_v g_{\overline{n},v}(t)$ will record the distribution of interest.

Let e_v be the standard basis vector with a 1 in coordinate v and 0s in all other coordinates. We derive the recurrence by considering the second-to-last card:

$$
\begin{aligned}
g_{\overline{m},v}(t) &= \sum_{\substack{D \in \mathcal{O}(\overline{m}) \\ D \dashv v}} t^{\theta(D)} = \sum_u \sum_{\substack{D \in \mathcal{O}(\overline{m}) \\ D \dashv uv}} t^{\theta(D)} \\
&= \sum_u \sum_{\substack{\tilde{D} \in \mathcal{O}(\overline{m}-e_v) \\ \tilde{D} \dashv u}} t^{\theta(\tilde{D}) + \frac{Z(D',u,v)}{n_u n_v}} = \sum_u t^{\frac{Z(D',u,v)}{n_u n_v}} g_{\overline{m}-e_v,u}.
\end{aligned}
\tag{13}
$$

This enables us to find $\overline{\kappa}_1(D')$ by recursively computing $g_{\overline{m},v}(t)$ for each card value v and for each integer vector $\overline{m}$ with $(0, \ldots, 0) \leq \overline{m} \leq (n_1, \ldots, n_k)$. There are $k \prod (n_i + 1)$ generating functions to compute, which is feasible for the cases we discuss.

6.1. Straight Poker. Straight poker is a game in which players receive 5 cards each from a deck of 52 distinct cards. The remaining cards are unused. Dealing is traditionally cyclic, so with 4 players the normal deal sequence is

$$D'_{\text{poker}} = (1234)^5 5^{32},$$

where 1, 2, 3, and 4 represent the players and 5 is the "hand" of unused cards. The reader may check that $Z(D'_{\text{poker}}, u, v)$ is the row u, column v entry of

$$Z(D'_{\text{poker}}) = \begin{pmatrix} 0 & 5 & 5 & 5 & 160 \\ -5 & 0 & 5 & 5 & 160 \\ -5 & -5 & 0 & 5 & 160 \\ -5 & -5 & -5 & 0 & 160 \\ -160 & -160 & -160 & -160 & 0 \end{pmatrix},$$

which allows us to use (13) to calculate

$$\overline{\kappa}_1(D'_{\text{poker}}) = \frac{104153993012865427259 9}{123600572196960202344} \approx 8.427.$$

The usual procedure in poker, however, is for the dealer to shuffle the cards and then allow the player to his left to cut them—that is, move some number of cards from the top to the bottom—before dealing. Fulman [**11**] showed that with a deck of distinct cards, a shuffle followed by a random cut was no more effective a randomizer than the shuffle alone.

When the deck will be dealt into hands, however, choosing a particular cut *can* enhance the randomness. Moving k cards has the same effect as making the deal sequence $\sigma^k D'_{\text{poker}}$, where σ is the cycle $(1, 2, \ldots, 52)$. The problem is small enough that we can simply try all possible cuts and report $\overline{\kappa}_1$ for each; the result is in Figure 3. The best place to cut the deck is after the 16th card, making $Z(u, 5) = 0$ for each player u. We find

$$\overline{\kappa}_1(\sigma^{16} D'_{\text{poker}}) = \frac{52348561969974736603 3}{126685078454994859800} \approx 4.132.$$

Thus a good cut can effectively halve the value of $\overline{\kappa}_1$, meaning it is worth one extra GSR shuffle. The method of the next section can be used to improve the situation further.

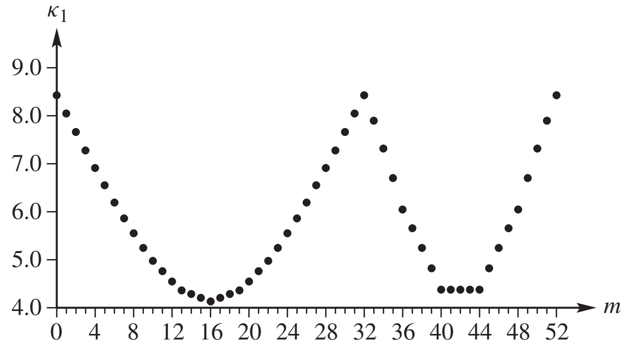


Figure 3. The effect on $\overline{\kappa}_1$ of cutting a poker deck at position m.

6.2. Bridge. As described in Section 3, a method of dealing bridge can be identified with a target deck $D' \in \mathcal{O}(\mathtt{N}^{13}\mathtt{E}^{13}\mathtt{S}^{13}\mathtt{W}^{13})$. We have $n = 52$ and $n_v = 13$ for each card value v, so

$$\overline{\kappa}_1(D') = \frac{1}{13N} \sum_{D \in \mathcal{O}(D')} \left| \sum_i Z(D', D(i), D(i+1)) \right|. \tag{14}$$

Suppose we deal a game of brige by "cutting the deck into hands." That is, the top 13 cards go to North, the next 13 to East, etc. Call this "ordered dealing." Symbolically,

$$D'_{\text{ord}} = \mathtt{N}^{13}\mathtt{E}^{13}\mathtt{S}^{13}\mathtt{W}^{13}.$$

There are 169 $\mathtt{N}$-$\mathtt{E}$ pairs in D'_{ord} and no $\mathtt{E}$-$\mathtt{N}$ pairs, so $Z(D'_{\text{ord}}, \mathtt{N}, \mathtt{E}) = 169$. Likewise for the other card values, so

$$Z(D'_{\text{ord}}) = 169 \begin{pmatrix} 0 & 1 & 1 & 1 \\ -1 & 0 & 1 & 1 \\ -1 & -1 & 0 & 1 \\ -1 & -1 & -1 & 0 \end{pmatrix},$$

where we give the values the implicit ordering $\mathtt{N} < \mathtt{E} < \mathtt{S} < \mathtt{W}$ and interpret the entries in the matrix accordingly. Using (13) we can compute

$$\overline{\kappa}_1(D'_{\text{ord}}) = \frac{93574839271687495932003418573}{33527961103430495524523400000} \approx 27.91.$$

Thus in the long run only slightly less shuffling is required for a bridge deck that will be cut into hands than for a deck in which all orderings are distinct.

Of course the way that most bridge players deal is cyclically:

$$D' = D'_{\text{cyc}} = (\mathtt{NESW})^{13}.$$

In that case the reader can check that

$$Z(D'_{\text{cyc}}) = 13 \begin{pmatrix} 0 & 1 & 1 & 1 \\ -1 & 0 & 1 & 1 \\ -1 & -1 & 0 & 1 \\ -1 & -1 & -1 & 0 \end{pmatrix},$$

which is to say, $Z(D'_{\text{cyc}}, u, v) = \frac{1}{13} Z(D'_{\text{ord}}, u, v)$ for all pairs of card types. It follows then from (14) that

$$\overline{\kappa}_1(D'_{\text{cyc}}) = \frac{1}{13}\overline{\kappa}_1(D'_{\text{ord}}) = \frac{7198064559360576610154109121}{3352796110343049552452340000} \approx 2.147.$$

So dealing cyclically works 13 times as well as simply cutting the deck into hands. That is to say, in the long run, switching from ordered to cyclic dealing is worth an extra $\log_2(13) \approx 3.7$ 2-shuffles.

The reason cyclic dealing is so much better than cutting into hands is that it makes $Z(D, u, v)$ small by better balancing the number of u-v pairs with the number of v-u pairs, for all u and v. If a is considerably larger than the deck size n, it is likely that when the deck is partitioned into a packets, all of the packets will either be empty or contain exactly one card. If such is the case, they will be riffled together in an arbitrary order, and the deck will be perfectly randomized. In fact, that suggests an alternate way to estimate how many shuffles are needed to adequately randomize a deck, an idea due to Reeds [**17**] and reported in Diaconis [**9**]. Simply calculate the likelihood that each card is in a different packet after the cut (this is the celebrated "birthday problem" of combinatorics), and pick a large enough that the probability is high.

The new idea here is that a dealing method can help ameliorate the bias in the case where a is still reasonably large, but some packet contains two cards. Imagine that the cards are initially arranged from "best" to "worst" before shuffling. Then a packet with two cards contains a "good" card atop a "worse" one, and after riffling the two cards will remain in the same order, though other cards may come between them. This is the source of the bias which remains even when a is large. If the two cards are dealt to players u and v, then we would like it to be approximately equally likely that u gets the good card and v the bad as the other way around. Thus we would like there to be about as many u-v pairs in the dealing method as there are v-u pairs.

Consider just the North and East players in bridge. We can describe a dealing method D' as it applies to those players by drawing a north-east lattice path starting from the lower-left corner of a 13×13 grid. That is, traverse D' and draw a north segment whenever an `N` is encountered and an east segment whenever an `E` is encountered. (The traditional names of players are very fortuitous for this exercise.) Every square to the southeast of the path has a northward segment to its left and an eastward segment above it, so it corresponds to a `N`-`E` pair in D'. Likewise the squares in the Young shape to the northwest of the path represent `E`-`N` pairs.

Figure 4 shows the paths and shapes for ordered and cyclic dealing on the left and in the middle. Cyclic dealing is much better than ordered dealing because the path stays near the diagonal of the grid, so about half the squares are on either side.

However, it always stays *to one side* of the diagonal, and thus it is easy to see that we can do better! The path on the right side of Figure 4 corresponds to

$$D'_{\text{bf}} = (\texttt{NESWWSEN})^6\texttt{NESW},$$

and by crossing the diagonal it balances the two sides as well as can be done, making $Z(D'_{\text{bf}}, \texttt{N}, \texttt{E}) = 1$. Likewise for the other pairs of players, so we have

$$Z(D'_{\text{bf}}) = \begin{pmatrix} 0 & 1 & 1 & 1 \\ -1 & 0 & 1 & 1 \\ -1 & -1 & 0 & 1 \\ -1 & -1 & -1 & 0 \end{pmatrix}$$

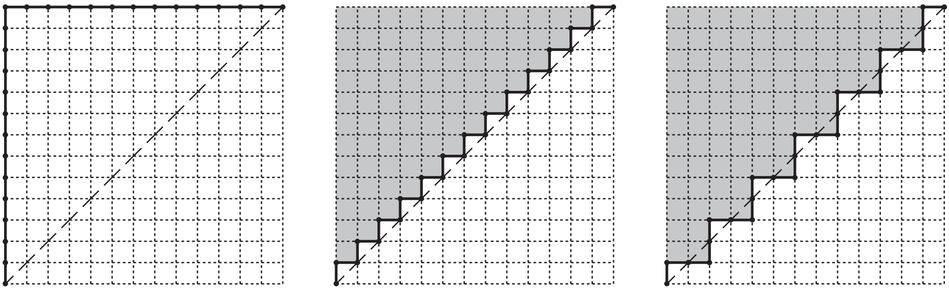


Figure 4. Three styles of bridge dealing, represented by lattice paths. The grid on the left represents ordered dealing, the center is cyclic dealing, and the one on the right is back-and-forth dealing. Each grid shows the sequence of `N` and `E` cards as north and east line segments respectively. The size of the Young shape to the northwest of the path is the number of `E-N` pairs in the target deck, and the size of the complementary shape is the number of `N-E` pairs.

and therefore

$$\overline{\kappa}_1(D'_{\mathrm{bf}}) = \frac{1}{13}\overline{\kappa}_1(D'_{\mathrm{cyc}}) = \frac{71980645593605766101541091 21}{435863494344596441818804200 00} \approx 0.165.$$

The dealing method described by D'_{bf} may be called "back-and-forth" dealing, since the dealer hands out cards once around the table clockwise, then once counterclockwise, then clockwise, counterclockwise, etc. We have shown that in the long run (i.e., for large a), back-and-forth dealing is 13 times as effective as cyclic dealing. Or, switching from cyclic to back-and-forth dealing is worth $\log_2(13) \approx 3.7$ extra GSR shuffles.

We could apply the same strategy to the poker deck that was cut after the 16th card; the combination of that cut and back-and-forth dealing produces a $\overline{\kappa}_1$ which is $1/5$ what it was with cyclic dealing.

7. TWO BIG QUESTIONS.

1. **How big does *a* have to be for the first-order approximation to be a good one?**

   By manipulating absolute value signs, one can show [**6**] that the error between variation distance from uniform and the first order estimate $\kappa_1 a^{-1}$ or $\overline{\kappa}_1 a^{-1}$ is bounded above by

$$\frac{1}{2}\sum_d \left\langle {n \atop d} \right\rangle \left| \frac{1}{a^n}\binom{a+n-d-1}{n} - \frac{1}{n!} - \frac{1}{a(n-1)!}\left(\frac{n-1}{2} - d\right) \right|.$$

   This bound, which depends only on the size of the deck (n) and the size of the shuffle (a), could undoubtedly be improved. Figure 2 shows that in the case of a deck of 52 distinct cards, the first-order estimate becomes quite good at about the point of the cutoff.

   So, how many times does a bridge player need to shuffle in order to see the promised benefits of switching dealing methods? Monte Carlo estimates (see [**6**] and [**8**] for a full explanation of methods and confidences) give strong evidence that back-and-forth dealing beats cyclic dealing after 5 or more shuffles. Table 1 shows the results.

Table 1. Variation distances from uniform after an a-shuffle for a deck of distinct cards and 3 methods of dealing bridge; $a = 16$ means 4 riffle suffles, $a = 32$ is 5 riffle shuffles, etc.

Deck	Method	$a = 16$	32	64	128	256	512	1024
52 Distinct $123 \cdots (52)$	Exact $44.0571a^{-1}$	1.0000 2.7536	0.9237 1.3768	0.6135 0.6884	0.3341 0.3442	0.1672 0.1721	0.0854 0.0860	0.0429 0.0430
Ordered Bridge $\mathtt{N}^{13}\mathtt{E}^{13}\mathtt{S}^{13}\mathtt{W}^{13}$	Monte Carlo $27.9095a^{-1}$	0.9902 1.7443	0.7477 0.8722	0.4230 0.4361	0.2183 0.2180	0.1104 0.1090	0.0550 0.0545	0.0274 0.0273
Cyclic Bridge $(\mathtt{NESW})^{13}$	Monte Carlo $2.1469a^{-1}$	0.2349 0.1342	0.0735 0.0671	0.0346 0.0335	0.0169 0.0168	0.0084 0.0084	0.0042 0.0042	0.0021 0.0021
Back-Forth Bridge $(\mathtt{NESWWSEN})^6(\mathtt{NESW})$	Monte Carlo $0.1651a^{-1}$	0.3118 0.0103	0.0260 0.0052	0.0073 0.0026	0.0022 0.0013	0.0008 0.0006	0.0003 0.0003	0.0002 0.0002

2. How good is the GSR model?

The GSR model represents idealized riffle shuffling, in the sense that every possible shuffle is equally likely. Human shufflers vary in their skill and "neatness," sometimes clumping cards together too much, sometimes not enough. The question is whether the conclusions drawn here from the GSR model will still hold when the model is replaced with the way real people shuffle cards. This is a topic for future work.

ACKNOWLEDGMENTS. The authors would like to thank the referees, and also Divakar Viswanath and Jeffrey Lagarias, for many helpful conversations.

REFERENCES

1. D. Aldous, Random walks on finite groups and rapidly mixing Markov chains, in *Seminar on Probability, XVII*, Lecture Notes in Mathematics, vol. 986, Springer, Berlin, 1983, 243–297.
2. S. Assaf, P. Diaconis, and K. Soundararajan, A rule of thumb for riffle shuffling, *Advances in Applied Probability* (to appear).
3. D. Bayer and P. Diaconis, Trailing the dovetail shuffle to its lair, *Ann. Appl. Probab.* **2** (1992) 294–313. doi:10.1214/aoap/1177005705
4. L. Carlitz, Eulerian numbers and polynomials, *Math. Mag.* **32** (1958/1959) 247–260.
5. M. Ciucu, No-feedback card guessing for dovetail shuffles, *Ann. Appl. Probab.* **8** (1998) 1251–1269. doi:10.1214/aoap/1028903379
6. M. Conger, Shuffling decks with repeated card values, Ph.D. dissertation, University of Michigan, Ann Arbor, MI, 2007.
7. M. Conger and D. Viswanath, Riffle shuffles of decks with repeated cards, *Ann. Probab.* **34** (2006) 804–819. doi:10.1214/009117905000000675
8. ———, Shuffling cards for blackjack, bridge, and other card games (2006), available at http://arxiv.org/abs/math/0606031v1.
9. P. Diaconis, *Group Representations in Probability and Statistics*, Institute of Mathematical Statistics Lecture Notes—Monograph Series, 11, Institute of Mathematical Statistics, Hayward, CA, 1988.
10. ———, Mathematical developments from the analysis of riffle shuffling, in *Groups, Combinatorics and Geometry (Durham, 2001)*, A. A. Ivanov, M. W. Liebeck, and J. Saxl, eds., World Scientific, River Edge, NJ, 2003, 73–97.
11. J. Fulman, Affine shuffles, shuffles with cuts, the Whitehouse module, and patience sorting, *J. Algebra* **231** (2000) 614–639. doi:10.1006/jabr.2000.8339
12. E. N. Gilbert, Theory of shuffling, Tech. Report MM-55-114-44, Bell Telephone Laboratories, New York, October 21, 1955.
13. R. L. Graham, D. E. Knuth, and O. Patashnik, *Concrete Mathematics*, 2nd ed., Addison-Wesley, Reading, MA, 1994.
14. B. Mann, How many times should you shuffle a deck of cards? in *Topics in Contemporary Probability and Its Applications*, Probability and Stochastics Series, CRC, Boca Raton, FL, 1995, 261–289.

15. A. A. Markov, Extension of the law of large numbers to dependent events, *Bull. Soc. Phys. Math.* **15** (1906) 135–156.
16. H. Poincaré, *Calcul des Probabilités*, deuxième ed., Gauthier-Villars, Paris, 1912.
17. J. Reeds, Unpublished manuscript, 1981.
18. S. Tanny, A probabilistic interpretation of Eulerian numbers, *Duke Math. J.* **40** (1973) 717–722. doi:10.1215/S0012-7094-73-04065-9
19. L. N. Trefethen and L. M. Trefethen, How many shuffles to randomize a deck of cards? *Proc. Roy. Soc. London Ser. A* **456** (2000) 2561–2568. doi:10.1098/rspa.2000.0625

MARK CONGER received his B.A. from Williams College in 1989 and his Ph.D. from the University of Michigan in 2007. In between he worked as a professional programmer for many years. He enjoys woodworking and taking things apart. He currently teaches at the University of Michigan.
Department of Mathematics, University of Michigan, 2074 East Hall, 530 Church Street, Ann Arbor, MI 48109
mconger@umich.edu

JASON HOWALD received his B.A. from Miami University in 1995 and his Ph.D. from the University of Michigan in 2001. He enjoys juggling, baking, and computer programming. He currently teaches at the State University of New York at Potsdam.
Department of Mathematics, SUNY Potsdam, 44 Pierpont Avenue, Potsdam, NY 13676
howaldja@potsdam.edu

. . . Mathematics Is . . .

“Somebody once said that philosophy is the misuse of a terminology which was invented just for this purpose. In the same vein, I would say that mathematics is the science of skillful operations with concepts and rules invented just for this purpose.”

Eugene Wigner, The unreasonable effectiveness
of mathematics in the natural sciences,
Comm. Pure Appl. Math. **13** (1960) 2.

—Submitted by Carl C. Gaither, Killeen, TX