The Annals of Applied Probability
1992, Vol. 2, No. 2, 294–313

TRAILING THE DOVETAIL SHUFFLE TO ITS LAIR

BY DAVE BAYER[1] AND PERSI DIACONIS[2]

Columbia University and Harvard University

We analyze the most commonly used method for shuffling cards. The main result is a simple expression for the chance of any arrangement after any number of shuffles. This is used to give sharp bounds on the approach to randomness: $\frac{3}{2}\log_2 n + \theta$ shuffles are necessary and sufficient to mix up n cards.

Key ingredients are the analysis of a card trick and the determination of the idempotents of a natural commutative subalgebra in the symmetric group algebra.

1. Introduction. The dovetail, or riffle shuffle is the most commonly used method of shuffling cards. Roughly, a deck of cards is cut about in half and then the two halves are riffled together. Figure 1 gives an example of a riffle shuffle for a deck of 13 cards.

A mathematically precise model of shuffling was introduced by Gilbert and Shannon [see Gilbert (1955)] and independently by Reeds (1981). A deck of n cards is cut into two portions according to a binomial distribution; thus, the chance that k cards are cut off is $\binom{n}{k}/2^n$ for $0 \le k \le n$. The two packets are then riffled together in such a way that cards drop from the left or right heaps with probability proportional to the number of cards in each heap. Thus, if there are A and B cards remaining in the left and right heaps, then the chance that the next card will drop from the left heap is $A/(A + B)$. Such shuffles are easily described backwards: Each card has an equal and independent chance of being pulled back into the left or right heap. An inverse riffle shuffle is illustrated in Figure 2.

Experiments reported in Diaconis (1988) show that the Gilbert–Shannon–Reeds (GSR) model is a good description of the way real people shuffle real cards. It is natural to ask how many times a deck must be shuffled to mix it up. In Section 3 we prove:

THEOREM 1. *If n cards are shuffled m times, then the chance that the deck is in arrangement π is $\binom{2^m + n - r}{n}/2^{mn}$, where r is the number of rising sequences in π.*

Rising sequences are defined and illustrated in Section 2 through the analysis of a card trick. Section 3 develops several equivalent interpretations of

Received January 1990; revised May 1991.
[1]Partially supported by the Alfred P. Sloan Foundation, by ONR contract N00014-87-K0214 and by NSF Grant DMS-90-06116.
[2]Partially supported by NSF Grant DMS-89-05874.
AMS 1980 *subject classifications*. 20B30, 60B15, 60C05, 60F99.
Key words and phrases. Card shuffling, symmetric group algebra, total variation distance.

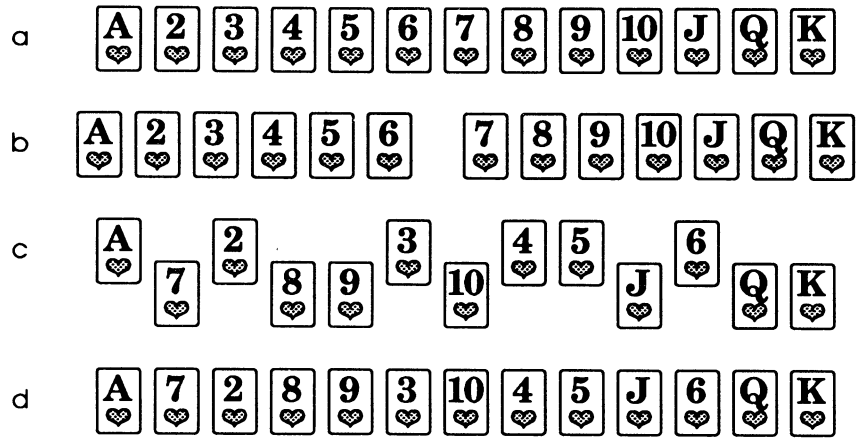

FIG. 1. *A riffle shuffle.* (a) *We begin with an ordered deck.* (b) *The deck is divided into two packets of similar size.* (c) *The two packets are riffled together.* (d) *The two packets can still be identified in the shuffled deck as two distinct "rising sequences" of face values.*

the GSR distribution for riffle shuffles, including a geometric description as the motion of n points dropped at random into the unit interval under the baker's transformation $x \to 2x \pmod 1$. This leads to a proof of Theorem 1.

Section 3 also relates shuffling to some developments in algebra. A permutation π has a descent at i if $\pi(i) > \pi(i+1)$. A permutation π has r rising sequences if and only if π^{-1} has $r-1$ descents. Let

$$A_k = \sum_{\pi \text{ has } k \text{ descents}} \pi$$

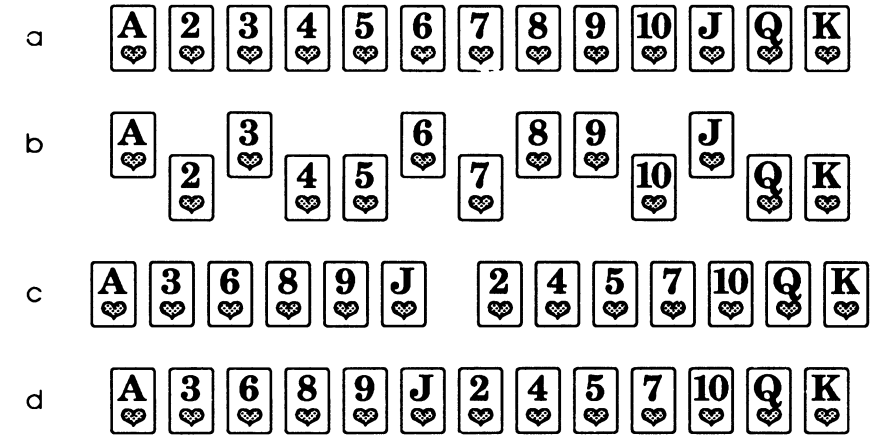

FIG. 2. *An inverse riffle shuffle.* (a) *We begin with a sorted deck.* (b) *Each card is moved one way or the other uniformly at random, to "pull apart" a riffle shuffle and retrieve two packets.* (c) *The two packets are placed in sequence.* (d) *The two packets can still be identified in the shuffled deck; they are separated by a "descent" in the face values. This shuffle is inverse to the shuffle diagrammed in Figure* 1.

TABLE 1
Total variation distance for m shuffles of 52 cards

m	1	2	3	4	5	6	7	8	9	10
$\|Q^m - U\|$	1.000	1.000	1.000	1.000	0.924	0.614	0.334	0.167	0.085	0.043

be defined as a formal linear combination in the group algebra. Following work by Solomon (1976), Gessel (1988) has shown that the A_k generate a commutative algebra. Theorem 1 gives a novel proof for an explicit expression of the primitive idempotents of this algebra, showing that it is generated by a formal sum corresponding to a GSR shuffle. This is related to recent work by Garsia and Reutenauer (1989) on free lie algebras, by Gerstenhaber–Schack (1987), Loday (1989) and Hanlon (1990) on Hodge decompositions of Hochschild homology.

In Section 4 we derive approximations to the total variation distance between the distribution after m shuffles and the uniform distribution. Here, if S_n is the symmetric group, U the uniform probability [so $U(\pi) = 1/n!$] and Q^m the Gilbert–Shannon–Reeds probability after m shuffles, then the total variation distance is defined as

$$\|Q^m - U\| = \max_{A \subset S_n} |Q^n(A) - U(A)|,$$

where, for example, $U(A) = \Sigma_{\pi \in A} U(\pi)$. Table 1 gives the total variation distance for 52 cards.

Table 1 shows that the total variation distance stays essentially at its maximum of 1 up to 5 shuffles, when it begins to decrease sharply by factors of 2 each time. This is an example of the cutoff phenomena described by Aldous and Diaconis (1986). Our analysis permits a sharp quantification of this:

THEOREM 2. *If n cards are shuffled m times with $m = \frac{3}{2}\log_2 n + \theta$, then for large n,*

$$\|Q^m - U\| = 1 - 2\Phi\left(\frac{-2^{-\theta}}{4\sqrt{3}}\right) + O\left(\frac{1}{n^{1/4}}\right),$$

with

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}\, dt.$$

Thus, the variation distance tends to 1 with θ small and to 0 with θ large.

A partial version of Theorem 2 has been proved by very different arguments in Aldous (1983). Thorp (1973) gives further references to the analysis of shuffling.

2. A card trick. Rising sequences, the basic invariant of riffle shuffling, were discovered by magicians Williams and Jordan at the beginning of this

century. A rising sequence is a maximal subset of an arrangement of cards, consisting of successive face values displayed in order. Rising sequences do not intersect, so each arrangement of a deck of cards is uniquely the union of its rising sequences. For example, the arrangement A, 5, 2, 3, 6, 7, 4 consists of the two rising sequences A, 2, 3, 4, and 5, 6, 7, interleaved together.

Suppose that a deck consists of n cards, arranged $1, 2, \ldots, n$. If a riffle shuffle divides the deck into packets of k and $n - k$ cards, then riffling together these packets interleaves cards $1, 2, \ldots, k$ with cards $k + 1, \ldots, n$. This creates two rising sequences: Cards $1, 2, \ldots, k$ remain in relative order within the deck, as do cards $k + 1, k + 2, \ldots, n$. Successive shuffles tend to double the number of rising sequences (until the capacity of the deck is approached), so shuffling a 52 card deck three times usually creates eight rising sequences. From these eight rising sequences, one can reconstruct exactly how the deck was shuffled.

This analysis of shuffling is the basis for a terrific card trick called "Premo" by Jordan. The performer removes a deck of cards from its case, hands it to a spectator and turns away from the spectators: "Give the deck a cut and a riffle shuffle. Give it another cut and another shuffle. Give it a final cut. I'm sure you'll agree that no living human could know the name of the top card. Remove this card, note its value, and insert it into the pack. Give the pack a further cut, a final shuffle, and a final cut."

Now the performer takes back the pack, spreads it in a wide arc on the table, and, after staring intensely, names the selected card.

To explain, consider what happens instead if the deck is never cut and the card is moved after the final shuffle. After three shuffles, the deck will usually have eight rising sequences, each consisting of an average of six and-a-half cards. Moving a card from the top to the middle of the deck usually creates a ninth rising sequence consisting of only the moved card, which is easily spotted.

What effect do cuts have on this analysis? Cuts respect the cyclic order of a deck, where card 1 follows card n. If we imagine the deck to be arranged in a loop, then cutting the deck rotates the loop. Thus, by allowing a spectator to cut the deck as often as desired, the performer merely gives up knowing where the loop starts. A riffle shuffle doubles this loop onto itself, analogous to the way that squaring doubles the unit circle in the complex plane onto itself. A search for successive face values in cyclic order winds once through an unshuffled deck and twice through a once shuffled deck. Depending on where the deck is taken to begin, this winding sequence could break up into two or three rising sequences.

If we view both the positions and face values of cards as having a cyclic order, then we can graph arrangements of cards on a torus, viewed as the product of two cycles. An unshuffled deck embeds as a (1, 1)-cycle and a once shuffled deck embeds as a (2, 1)-cycle. One sees that rising sequences are an artifact of where the torus is cut to make a square.

Define the *winding number* of a deck to be the number of laps required to cycle through the deck by successive face values. A deck begins with winding

number 1; each of the first few shuffles of a deck doubles its winding number. Moving a card usually increases the winding number by 1. We can identify the moved card by associating a count with each card, giving the total number of cards between its predecessor and successor, as we follow the winding sequence through the deck: Let $\sigma(i)$ give the position of card i, and let $d(i, j)$ be the least positive integer so $d(i, j) = \sigma(j) - \sigma(i) \pmod{n}$. Then we associate with each card i the count $d(i-1, i) + d(i, i+1) - 1$. Ideally, the moved card will sit on its own lap of the winding sequence and its count will be the only count greater than n.

The trick as described is not sure-fire. To investigate, we performed various Monte Carlo experiments. As expected, the trick is most successful when the card is moved after the final shuffle. We programmed a computer to shuffle the cards m times according to the GSR distribution, cut the deck uniformly at random, move the top card to a binomially distributed position and then cut the deck again. From here, the computer made and scored its guesses as to which card was moved, using the strategy described above: Given k guesses, it chose the k cards with highest counts, breaking ties at random (see Table 2).

With three shuffles, the trick succeeded in 84% of 1,000,000 trials. With two guesses allowed, the success rate went up to 94%. This is a reasonable rate to aspire to in practice; if the performer suspects two cards, a leading question like "your card was a red card" will resolve the ambiguity. Table 2 gives results for m shuffles, $2 \le m \le 12$, and k guesses at the card, $k = 1, 2, 3, 13, 26$.

In studying these numbers, we were most struck by the results for many shuffles. Already at four shuffles, this trick is terrible magic, but even at eight shuffles, it can still make a great bet: Betting even money on being able to pick the moved card with 26 guesses, one enjoys nearly a 10% advantage. This is startling, considering that people rarely shuffle eight times in practice.

Observe also that the advantage halves after each shuffle, in the limit. Trying to explain this effect lead us to the results in the remainder of this paper.

We conclude this section with a brief history of the magical use of riffle shuffles. The earliest clear application of rising sequences that we know of is due to C. O. Williams, a respected inventor of magic who worked at the turn of

TABLE 2

Probability of success in Jordan's card trick with 52 cards shuffled m times, and 1, 2, 3, 13 or 26 guesses allowed. Each entry is based on 1,000,000 Monte Carlo trials. All entries are given in thousandths

m	2	3	4	5	6	7	8	9	10	11	12	∞
1	997	839	288	088	042	028	023	021	020	020	019	019
2	1000	943	471	168	083	057	047	042	040	039	039	038
3	1000	965	590	238	123	085	070	063	061	059	058	058
13	1000	998	884	617	427	334	290	270	260	254	252	250
26	1000	999	975	835	688	596	548	524	513	505	503	500

the century. Williams allowed a prearranged deck to be shuffled once. The deck also had its backs aligned all in the same direction. When shuffled, one half was turned end for end so the two rising sequences could be clearly identified and a "card reading" was performed. See Williams (1912).

A much more sophisticated set of applications was invented by the American Jordan, who was an inventor of magic, designer of radios, professional contest winner and chicken farmer. The first mention of his work appears in Jordan (1916). This contains a description of a trick called "long distance mindreading." In effect, "you mail an ordinary pack of cards to anyone, requesting them to shuffle and select a card. He shuffles again and returns only half the pack to you, not intimating whether or not it contains his card. By return mail you name the card he selected."

Further information on the mathematics of shuffling can be found in Gardner (1966, 1977).

3. Shuffles and their generalizations. The Gilbert–Shannon–Reeds model for shuffling has alternate descriptions, and a natural generalization to shuffles that begin with the deck being cut into a packets, with $a \geq 2$; the various packets are then riffled together.

GEOMETRIC DESCRIPTION. The geometric model begins by placing n points uniformly and independently in the unit interval. The points are labeled in the order $x_1 < x_2 < \cdots < x_n$. For positive integral a, the map $x \to ax \pmod 1$ maps $[0, 1]$ onto itself and preserves measure. This map rearranges the points x_i and so gives a measure on the symmetric group which will be called an a-shuffle.

A 2-shuffle is like an ordinary riffle shuffle: Points in $[0, \frac{1}{2}]$ and $[\frac{1}{2}, 1]$ are stretched out and interlaced.

MAXIMUM ENTROPY DESCRIPTION. All possible ways of cutting a deck into a packets and then interleaving the packets are equally likely. Empty packets are allowed.

INVERSE DESCRIPTION. All possible ways of pulling a shuffled deck back apart into a packets are equally likely. Empty packets are allowed.

The following generates an inverse a-shuffle with the correct probability: A deck of n cards is held face down. Successive cards are turned face up and dealt into one of a piles uniformly and independently. After all cards have been distributed, the piles are assembled from left to right and the deck is turned face down.

An example of an inverse 2-shuffle was illustrated in Figure 2.

SEQUENTIAL DESCRIPTION. Choose integers $j_1, j_2, \ldots, j_a$ according to the multinomial distribution

$$P(j_1, \ldots, j_a) = \binom{n}{j_1 \cdots j_a} \frac{1}{a^n}.$$

Thus, $0 \le j_i \le n$, $\sum_{i=1}^{a} j_i = n$ and the j_i have the same distribution as the number of balls in each box i if n balls are dropped at random into a boxes.

Given j_i, cut off the top j_1 cards, the next j_2 cards and so on, producing a or fewer packets. Shuffle the first two packets using the GSR shuffle described in Section 1. Then shuffle this combined packet with packet 3, and so forth. This is equivalent to riffling all a packets together at once, where if there are A_i cards remaining in each heap, the chance that the next card will drop from heap i is $A_i/(A_1 + \cdots + A_a)$.

Shuffles of this type are performed by casinos to mix several decks. For example, 104 cards are sometimes mixed by cutting into four piles, shuffling packets 1 and 2 together, then packets 3 and 4 together and finally 1 and 2 into 3 and 4. This is equivalent to a 4-shuffle.

LEMMA 1. *The four descriptions generate the same permutation distribution. Moreover, in each model an a-shuffle followed by a b-shuffle is equivalent to an ab-shuffle.*

PROOF. Each description results in a multinomial number of cards in each packet. This holds by decree for the sequential description and is clear for the inverse description. For the geometric description, the packet sizes are determined by how many points are chosen in each interval $[(i-1)/a, (i/a)]$, which is also multinomial. For the maximum entropy description, the number of possible interleavings starting from a given cut is multinomial; they are in 1 : 1 correspondence with the ways of dividing a deck into a subsets with the corresponding packet sizes.

Given the packet sizes, the maximum entropy description asserts that all possible interleavings are equally likely. This also clearly holds for the inverse description. For the sequential description, observe that when the first two piles of size j_1, j_2 are shuffled, the chance of any specific sequence of left-right drops is

$$\frac{j_1(j_1-1)\cdots 1 \cdot j_2(j_2-1)\cdots 1}{(j_1+j_2)(j_1+j_2-1)\cdots 1} = \binom{j_1+j_2}{j_1}^{-1}.$$

When these cards are shuffled into the third packet of size j_3, all $\binom{j_1+j_2+j_3}{j_3}$ positions for its cards are equally likely. This continues to hold for each successive packet.

No state information is retained between shuffles in these three models, so the product rule for a sequence of shuffles holds in each model once it is established for one. This easily follows from the inverse description: Lexicographically combining the pile assignments from an inverse a-shuffle and an inverse b-shuffle yields uniform and independent pile assignments for an inverse ab-shuffle.

For the geometric description, the lemma follows from the independence of base a digits of points picked uniformly in $[0, 1]$: Choosing n points in $[0, 1]$,

labeling them with their leading digits and applying the map $x \to ax \pmod 1$, is the same as choosing n points in $[0, 1]$ and labeling them arbitrarily with integers from $\{0, \ldots, a-1\}$. Thus, all interleavings are equally likely for a given set of packet sizes. Moreover, the sets of n points which yield a given shuffle map to sets of n points distributed uniformly in $[0, 1]$. Thus, successive shuffles can reuse the points x_i without first having to reposition them at random in $[0, 1]$, so the product rule follows from the identity

$$b(ax \pmod 1) \pmod 1 = abx \pmod 1. \qquad \square$$

The main result of this section gives an explicit description of a-shuffles. To state it, we need to specify how we are associating shuffles with permutations. If a shuffle transforms an arrangement $1, 2, 3, 4, 5$ of cards into the arrangement $2, 3, 4, 5, 1$, then we associate this shuffle with the permutation

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 4 & 5 & 1 \end{pmatrix}.$$

THEOREM 3. *The probability that an a-shuffle will result in the permutation π is*

$$\frac{\binom{a+n-r}{n}}{a^n}, \quad \textit{where } r \textit{ is the number of rising sequences in } \pi.$$

PROOF. Using the maximum entropy description, this probability is determined by the number of ways of cutting an ordered deck into a packets, so π is a possible interleaving. Because each packet stays in order as the cards are riffled together, each rising sequence in the shuffled deck is a union of packets. Thus, we want to count the number of ways of refining r rising sequences into a packets.

We emulate the classical stars and bars argument, counting arrangements of cuts on the ordered deck before shuffling: At least one cut must fall between each successive pair of rising sequences of π, but the remaining cuts can be located arbitrarily. Thus, the n cards form dividers creating $n+1$ bins, into which the $a-r$ spare cuts are allocated. There are $\binom{a+n-r}{n}$ ways of doing this. There are a^n possible a-shuffles in all, giving the stated probability. $\square$

One could keep track of the packet structure of a shuffle by coloring the packets before they are riffled together. Then the above enumeration counts colorings of π which look like they came from a shuffle. This interpretation of the proof of Theorem 3 is illustrated in Figures 3 and 4.

We first proved this theorem via the geometric description of a-shuffles, viewing them as baker's transformations on the n-cube. With Laurie Beckett's help, we were able to isolate from this proof the purely combinatorial argument given above.

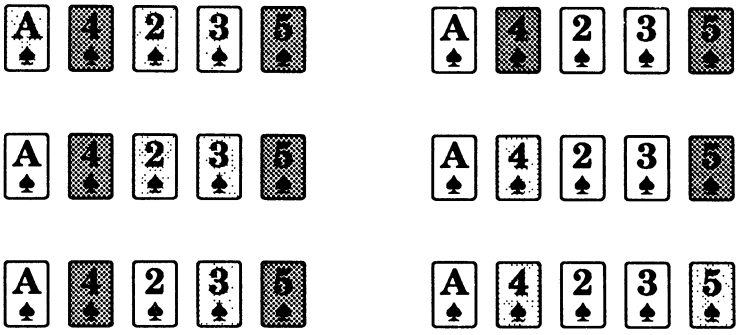

FIG. 3. *Six 3-colorings of the arrangement* A, 4, 2, 3, 5. *Dye the packets of a 3-shuffle a sequence of distinct colors, before the packets are riffled together. Because the relative order of the cards in each packet is preserved by shuffling, such colorings refine the decomposition of the deck into rising sequences. Each 3-coloring with this property arises from a 3-shuffle, so the probability of any arrangement is determined by the number of ways it can be so colored.*

Summing the formula of Theorem 3 over all permutations π gives 1. There are $A_{n,r}$ permutations with r rising sequences, where the $A_{n,r}$ are the Eulerian numbers. Thus, multiplying this sum by a^n gives

$$a^n = \sum_{r=1}^{n} A_{n,r}\binom{a+n-r}{n},$$

which is Worpitzky's identity. Conversely, a proof of Theorem 3 can be inferred from the proof of Theorem 4.5.14 in Stanley (1986), by considering the trivial poset consisting only of incomparable elements. Stanley's result is an extension of Worpitzky's identity to partially ordered sets, which may have interesting shuffling interpretations.

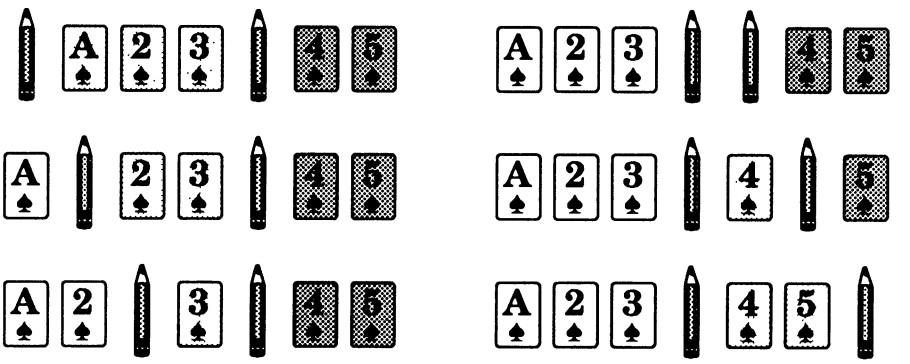

FIG. 4. *Counting the possible 3-colorings of* A, 4, 2, 3, 5. *Use two pencils to mark where the deck* A, 2, 3, 4, 5 *is going to be cut into packets for 3-shuffling and dye the cards accordingly. The arrangement* A, 4, 2, 3, 5 *consists of the two rising sequences* A, 2, 3 *and* 4, 5. *Cuts which can produce this arrangement are enumerated by placing one pencil between the two rising sequences and placing the other pencil arbitrarily.*

COROLLARY 1. *If a deck of cards is given a sequence of m shuffles of types* $a_1, a_2, \ldots, a_m$, *then the chance that the deck is in arrangement* π *is given by*

$$\frac{\binom{n+a-r}{n}}{a^n},$$

where $a = a_1, a_2 \cdots a_k$ *and* r *is the number of rising sequences in* π.

PROOF. Combine Lemma 1 with Theorem 3. □

Theorem 1 of the Introduction follows immediately as a corollary.

Theorem 3 generalizes earlier work of Shannon, who showed that if n cards are 2-shuffled m times, with $m \leq \log_2 n$, then all arrangements with 2^m rising sequences have the same probability. The corollary is used in Section 4 to give exact results and asymptotics for total variation convergence to the uniform measure.

The next corollary shows that the number of rising sequences forms a Markov chain. As motivation, note that a function of a Markov chain is not usually Markov. Heller (1965) gives a complicated necessary and sufficient condition and Rogers and Pitman (1981) give a sufficient condition which is easy to check in the present case.

COROLLARY 2. *Let a Markov chain on the symmetric group begin at the identity and proceed by successive independent a-shuffles chosen from the Gilbert–Shannon–Reeds measure. Then* $R(\pi)$, *the number of rising sequences, forms a Markov chain.*

PROOF. From Theorem 3, the conditional law of π given $R(\pi)$ is uniform. Rogers and Pitman [(1981), Lemma 1] show that this, coupled with a completeness condition on the induced family of distributions for the process of rising sequences, is sufficient. In the present setting, completeness amounts to showing

$$\sum_{i=1}^{m} \binom{a^m + n - r}{n} f(r) = \sum_{i=1}^{n} \binom{a^m + n - r}{n} g(r) \quad \text{for } m = 0, 1, 2, \ldots,$$

implies $f = g$. The left side is the polynomial

$$\frac{1}{n!}\big[(x+n-1)(x+n-2)\cdots xf(1)$$
$$+(x+n-2)(x+n-3)\cdots(x-1)f(2)$$
$$+\cdots+x(x-1)\cdots(x-(n-1))f(n)\big],$$

evaluated at $x = a^m$. Evaluating at $x = i$ gives $f(i)$. □

In unpublished work, Gessel has derived formulas for the transition matrix of the induced chain. Of course, the first row of the matrix is given by Theorem 3. The other rows are not as simple to write out.

The final corollary connects shuffling to results in algebra. To describe things, define the group algebra $L(S_n)$ as the set of all functions from S_n into the rational numbers Q. Elements of L may be thought of as formal linear combinations of permutations with rational coefficients. Multiplication is given by formally multiplying the linear expressions using the multiplication on S_n. This is the same as convolving together the associated functions. In $L(S_n)$ let

$$A_i = \sum_{R(\pi)=i} \pi, \qquad i = 1, 2, \ldots, n.$$

COROLLARY 3. *Let $\mathscr{A}$ be the subalgebra of $L(S_n)$ generated by $A_1, \ldots, A_n$. Then $\mathscr{A}$ is a commutative, semisimple algebra of dimension n. A basis of primitive idempotents is given by $e_n(l) = \sum_{r=1}^{n} \sigma_l(n-r, \ldots, 1-r)A_r$, with σ_l the lth elementary symmetric function.*

PROOF. Using the theorem, an a-shuffle can be represented in $\mathscr{A}$ as

$$B_a = \frac{1}{a^n} \sum_{r=1}^{n} \binom{a+n-r}{n} A_r.$$

Now B_2^2 is a positive linear combination of $A_1 = Id$, A_2 and A_2^2. Theorem 3 gives $B_2^2 = B_4 \in \mathscr{A}$. Thus A_2^2 is in $\mathscr{A}$. Next, $B_2 B_3$ is a positive linear combination of A_2A_3, A_2^2, A_2, A_3 and A_1. It follows that $A_2A_3 \in \mathscr{A}$ and, from $B_2B_3 = B_3B_2 = B_6$, $A_2A_3 = A_3A_2$. From here A_3^2 and then $A_iA_3 = A_3A_i$ are in $\mathscr{A}$. Continuing inductively proves that $\mathscr{A}$ is a commutative algebra.

To complete the proof, consider successive powers of B_2:

$$B_2^m = B_{2^m} = \frac{1}{2^{mn}} \sum_{r=1}^{n} \binom{2^m+n-r}{n} A_r$$

$$= \frac{1}{n!} \sum_{l=0}^{n-1} \frac{1}{2^{lm}} \sum_{r=1}^{n} \sigma_l(n-r, \ldots, 1-r) A_r.$$

From this it follows that the linear map $\mathscr{A} \to \mathscr{A}$ given by multiplying by B_2 has distinct eigenvalues $1, 1/2, 1/2^2, \ldots, 1/2^{n-1}$. It is thus diagonalizable with the $e(l)$ as eigenvectors.

Left multiplication on itself gives a faithful representation of $\mathscr{A}$ as a commutative matrix algebra which we have shown contains an element B_2 with distinct eigenvalues. It follows that the set of matrices that commute with B_2 is all polynomials in B_2. Thus B_2 generates $\mathscr{A}$ (since elements of $\mathscr{A}$ commute with B_2). Thus, the $e(l)$ simultaneously diagonalize $\mathscr{A}$ which is therefore semisimple. □

The algebra $\mathscr{A}$ has appeared in various areas of mathematics. We describe some of these briefly.

HOCHSCHILD COHOMOLOGY. The homomorphism $\operatorname{sgn}(\pi)$ extends to an automorphism of $L(S_n)$ by linearity: $\sum a_\pi \pi \mapsto \sum (-1)^\pi a_\pi \pi$. The image of $\mathscr{A}$ under sgn figures prominently in the work of Barr (1968), Gerstenhaber and Schack (1987), Loday (1989) and Hanlon (1990). These authors are concerned with Hodge-type decompositions of a complex made by tensoring commutative algebras with boundary maps typified by $a_1 \otimes a_2 \otimes a_3 \mapsto a_1 a_2 \otimes a_3 - a_1 \otimes a_2 a_3$. If $s_n = \operatorname{sgn}(A_2)$, they show that $\partial s_n = s_{n-1}\partial$ and split the usual Hochschild cohomology using this action. They show that the indempotents of the algebra $\mathscr{A}$ give a basis for all such decompositions. Fix an idempotent $e_n(j)$ and consider $e_n(j)L(S_n)$ as a representation of S_n. Hanlon (1990) determines the dimensions and other properties of these representations.

COXETER GROUPS. Solomon (1976) has introduced related algebras for general Coxeter groups. Specializing to the symmetric group S_n, say that a permutation π has a descent at i if $\pi(i) > \pi(i+1)$. This agrees with the definition we have been using for arrangements of decks of cards. Permutations can have descents at positions $1, 2, \ldots, n-1$. Observe that π has $r-1$ descents if and only if π^{-1} has r rising sequences: the ith entry of π gives the position of the letter i in π^{-1} and is a descent whenever $i+1$ begins a new rising sequence of π^{-1}. Thus, the algebra $\mathscr{A}$ studied in Corollary 3 could, up to isomorphism, have been defined in terms of descents. More generally, let $D(\pi)$ be the descent set of π. Define

$$A_S = \sum_{D(\pi)=S} \pi.$$

Solomon showed that A_S forms a noncommutative algebra with a natural geometric interpretation.

The word $A_{\{1\}}$ has a natural shuffle interpretation: remove a card at random and put it back on top. This generates a probability on S_n represented as $(1/n)(id + A_{\{1\}})$ in the group algebra. Aldous and Diaconis (1986) show that it takes $n \log n + cn$ iterations to get close to uniform. Diaconis and Pitman (1991) give arguments entirely analogous to the ones in the present paper, for a generalization of these shuffles.

Descents and rising sequences can be defined for any Coexter group. For example, the hyperoctahedral group B_n of symmetries of an n-dimensional cube can be represented as the group of all n by n signed permutation matrices. If we write these matrices as signed permutations and order the letters so $-n < \cdots -1 < 1 < \cdots < n$, then descents and rising sequences can be defined as usual. There is a card shuffling interpretation to the descent algebra: If a packet of k cards is cut off, flipped over and riffled face up into the remaining $n-k$ cards according to the GSR distribution, then the resulting distribution can be represented as a linear combination of the identity and group elements with one rising sequence. Section 5 develops this further.

4. The approach to uniformity. This section analyzes the approach to uniformity. We derive approximations when n is large after m shuffles, with $m = \frac{3}{2}\log_2 n + \theta$. For notational convenience, write $m = \log_2(n^{3/2}c)$, so $c = 2^\theta$ satisfies $0 < c < \infty$. The arguments use the asymptotics of Eulerian numbers. Asymptotics and exact results are compared at the end of this section.

We first develop a local limit theorem. This gives a tractable approximation for the probability that m shuffles will result in a given permutation.

PROPOSITION 1. *Let* $Q^m(r) = \binom{2^m + n - r}{n}/2^{mn}$ *be the probability of a permutation with* r *rising sequences after* m *shuffles from the GSR distribution. Let* $r = n/2 + h$, $-n/2 + 1 \le h \le n/2$. *Let* $m = \log_2(n^{3/2}c)$ *with* $0 < c < \infty$ *fixed. Then*

$$Q^m(r) = \frac{1}{n!}\exp\left\{\frac{1}{c\sqrt{n}}\left(-h + \frac{1}{2} + O_c\left(\frac{h}{n}\right)\right)\right. \\ \left. - \frac{1}{24c^2} - \frac{1}{2}\left(\frac{h}{cn}\right)^2 + O_c\left(\frac{1}{n}\right)\right\}. \tag{4.1}$$

PROOF.

$$Q^m(r) = \frac{1}{n!}\left(\frac{2^m + n - r}{2^m} \cdots \frac{2^m + 1 - r}{2^m}\right) \\ = \frac{1}{n!}\exp\left\{\sum_{i=0}^{n-1}\log\left(1 + \frac{(n/2) - h - i}{cn^{3/2}}\right)\right\}.$$

The logarithmic terms can be upper and lower bounded using

$$x - \frac{x^2}{2} + \frac{x^3}{3} - x^4 \le \log(1 + x) \le x - \frac{x^2}{2} + \frac{x^3}{3}, \qquad -\frac{1}{2} < x < 1.$$

Standard summation formulas give

$$\frac{1}{cn^{3/2}}\sum_{i=0}^{n-1}\left(\frac{n}{2} - h - i\right) = \frac{-h + 1/2}{c\sqrt{n}},$$

$$\frac{1}{2c^2n^3}\sum_{i=0}^{n-1}\left(\frac{n}{2} - h - i\right)^2 = \frac{1}{24c^2} + \frac{1}{2}\left(\frac{h}{cn}\right)^2 + O_c\left(\frac{1}{n}\right),$$

$$\frac{1}{3c^3n^{9/2}}\sum_{i=0}^{n-1}\left(\frac{n}{2} - h - i\right)^3 = O_c\left(\frac{h}{n^{3/2}}\right),$$

$$\frac{1}{c^4n^6}\sum_{i=0}^{n-1}\left(\frac{n}{2} - h - i\right)^4 = O_c\left(\frac{1}{n}\right). \qquad \square$$

The probabilities $Q^m(r)$ are monotone decreasing in r for fixed m. The next proposition determines when they cross the point $1/n!$. This is crucial for

analyzing total variation, because it determines the set on which the total variation is achieved.

PROPOSITION 2. *With notation as in Proposition* 1, *let* h^* *be an integer such that* $Q^m(n/2 + h) \geq 1/n! \Leftrightarrow h \leq h^*$. *Then, for any fixed* c, *as* $n \to \infty$,

$$h^* = \frac{-\sqrt{n}}{24c} + \frac{1}{12c^3} + B + O_c\left(\frac{1}{\sqrt{n}}\right), \tag{4.2}$$

where $-1 \leq B \leq 1$.

PROOF. $Q^m(n/2 + h) \geq 1/n!$ if and only if the exponent in (4.1) is nonnegative. Setting the exponent equal to 0 and collecting terms gives (4.2). □

The next result is the main theorem of this section, combining the estimates above to give the asymptotics of the total variation. The argument uses standard results about sums of independent variables; see Feller (1971) for background.

THEOREM 4. *Let* Q^m *be the Gilbert–Shannon–Reeds distribution on the symmetric group* S_n. *Let* U *be the uniform distribution. For* $m = \log_2(n^{3/2}c)$, *with* $0 < c < \infty$ *fixed, as* n *tends to* ∞,

$$\|Q^m - U\| = 1 - 2\Phi\left(\frac{-1}{4c\sqrt{3}}\right) + O_c\left(\frac{1}{n^{1/4}}\right)$$

with $\Phi(x) = \int_{-\infty}^{x} e^{-t^2/2}\, dt/\sqrt{2\pi}$.

PROOF. With notation as in Propositions 1 and 2 above, $\|Q^m - U\|$ equals

$$\sum_{-n/2 < h \leq h^*} R_{nh}\left(Q^m\left(\frac{n}{2} + h\right) - \frac{1}{n!}\right), \tag{4.3}$$

where R_{nh} is the number of permutations with $n/2 + h$ rising sequences. This uses the fact that the number of rising sequences is a sufficient statistic for both Q^m and U as explained in Section 3 and that total variation between two probabilities equals the total variation between the induced laws of any sufficient statistic [see, e.g., Diaconis and Zabell (1982), Lemma 6.1].

A permutation has r rising sequences if and only if π^{-1} has $r - 1$ descents; see Section 3 for further discussion. The number of permutations with j descents is called the Eulerian number a_{nj}; see Tanny (1973), Stanley (1977) and other papers in the latter volume. Tanny and Stanley show that $a_{nj}/n!$ equals the chance that the sum of n random variables uniform on [0, 1] is

between j and $j+1$. Thus $a_{nj}/n!$ and $R_{nh}/n!$ obey the central limit theorem as in Tanny (1973). In particular if $x_n = h/\sqrt{n/12}$, the local limit theorem gives

$$\frac{R_{nh}}{n!} = \frac{e^{-(1/2)x_n^2}}{\sqrt{2\pi n/12}}\left(1 + o\left(\frac{1}{\sqrt{n}}\right)\right) \quad \text{uniformly in } h. \tag{4.4}$$

The usual form of the central limit theorem for the distribution function of $a_{nj}/n!$ and $R_{nh}/n!$ gives

$$\frac{1}{n!}\sum_{h=-n/2}^{h^*} R_{nh} = \Phi\left(\frac{-1}{4c\sqrt{3}}\right)\left(1 + O\left(\frac{1}{\sqrt{n}}\right)\right) \quad \text{uniformly.}$$

The sum (4.3) can be broken into two zones. Recall from Proposition 2 that $h^* = -\sqrt{n}/24c + O(1)$:

$$\text{zone 1:} \quad \left\{\frac{-10n^{3/4}}{\sqrt{c}} \le h \le h^*\right\} = I_1,$$

$$\text{zone 2:} \quad \left\{-\frac{n}{2} \le h < \frac{-10n^{3/4}}{\sqrt{c}}\right\} = I_2.$$

As will be shown, only zone 1 contributes.

From (4.4) and Proposition 1, $\Sigma_{I_1} R_{nh} Q^m(n/2 + h)$ equals

$$\frac{e^{-1/24c^2}}{\sqrt{2\pi n/12}}\sum_{I_1}\exp\left\{-\frac{1}{2}\left(\frac{h}{\sqrt{n/12}}\right)^2 - \frac{h}{c\sqrt{n}} + O_c\left(\frac{1}{n^{1/4}}\right)\right\}\left\{1 + o\left(\frac{1}{\sqrt{n}}\right)\right\}$$

$$= \frac{e^{-1/24c^2}}{\sqrt{2\pi}}\int_{-\infty}^{-(2c\sqrt{12})^{-1}} e^{-x^2/2 - x/c\sqrt{3}}\,dx\left(1 + O\left(\frac{1}{n^{1/4}}\right)\right)$$

$$= \Phi\left(\frac{1}{4c\sqrt{3}}\right)\left(1 + O\left(\frac{1}{n^{1/4}}\right)\right).$$

In zone 2, $Q^m(n/2 + h) \le Q^m(1) \le e^{\sqrt{n}/2c}/n!$. The standard large deviations bound as in Feller [(1971), Chapter 16] applied to the sum of n uniforms shows

$$\sum_{I_2}\frac{R_{nh}}{n!} \sim \frac{1}{10n^{1/4}\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{10\sqrt{12}\,n^{1/4}}{\sqrt{c}}\right)^2\right].$$

Combining bounds completes the proof. □

TABLE 3
Total variation distance for m shuffles of 25, 32, 52, 78, 104, 208 *or* 312 *distinct cards*

m	1	2	3	4	5	6	7	8	9	10
25	1.000	1.000	0.999	0.775	0.437	0.231	0.114	0.056	0.028	0.014
32	1.000	1.000	1.000	0.929	0.597	0.322	0.164	0.084	0.042	0.021
52	1.000	1.000	1.000	1.000	0.924	0.614	0.334	0.167	0.085	0.043
78	1.000	1.000	1.000	1.000	1.000	0.893	0.571	0.307	0.153	0.078
104	1.000	1.000	1.000	1.000	1.000	0.988	0.772	0.454	0.237	0.119
208	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.914	0.603	0.329
312	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.883	0.565

Theorem 2 of the Introduction follows immediately as a corollary.

REMARK 1. The function $1 - 2\Phi(-1/4c\sqrt{3})$ has the following asymptotic behavior:

$$1 - 2\Phi\left(\frac{-1}{4c\sqrt{3}}\right) \sim \frac{1}{2c\sqrt{6\pi}} \quad \text{as } c \to \infty,$$

$$1 - 2\Phi\left(\frac{-1}{4c\sqrt{3}}\right) \sim 1 - \frac{4c\sqrt{3}}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{-1}{4c\sqrt{3}}\right)^2\right] \quad \text{as } c \to 0.$$

Note that $m = \log_2(n^{3/2}c)$ has c inside the logarithm, so $c = 2^j$ where j is the number of shuffles beyond $\frac{3}{2}\log_2 n$ that have been performed. It follows that the variation distance tends to 0 exponentially in j for j positive. It tends to 1 doubly exponentially in j for j negative.

REMARK 2. The asymptotics show that about $\frac{3}{2}\log_2 n$ shuffles are needed to mix up n cards. Table 3 gives exact computations of variation distance for a number of popular deck sizes. Each deck size shows the cutoff phenomenon. Variation distance decreases by a factor of 2 after each shuffle following the cutoff. For comparison, Table 4 gives $\frac{3}{2}\log_2 n$ for these deck sizes.

REMARK 3. The appearance of the normal distribution in the asymptotics for total variation is not an accident; see Diaconis, Graham and Morrison (1990). With m shuffles, the number of rising sequences is normally distributed as it is under the uniform distribution. Variation distance equals the distance between the two limiting normal distributions.

TABLE 4
Shuffles needed to mix 25, 32, 52, 78, 104, 208 *or* 312 *cards*

n	25	32	52	78	104	208	312
$\frac{3}{2}\log_2 n$	6.97	7.50	8.55	9.43	10.05	11.55	12.43

5. Three developments.

5.1. *A different distance.* The total variation distance, while standard in probability theory, can be difficult to explain to nonspecialists. The following alternative distance has proved useful and is of interest on its own.

Consider the following problem: A deck of cards is face down on the table. A guesser tries to guess at the cards one at a time. After each guess, the current top card is turned over to reveal its value and then discarded. If the guesser believes that the deck is well mixed, the optimal strategy is to guess any card first (chance $1/52$ of being correct) and thereafter guess a card known to be in the deck. The expected number of correct guesses is

$$\frac{1}{52} + \frac{1}{51} + \cdots + 1 = 4.54.$$

With n cards, the number correct is asymptotically normal with mean $\log n$ and variance $\log n$.

Suppose that the deck has been riffle shuffled k times, where k is unknown. The starting assignment is assumed known. We do not know the optimal strategy. A conjectural optimal strategy was used to produce Table 5.

The table shows the result of a Monte Carlo experiment. Each cell is based on 100,000 trials. A cell shows the average number of correct guesses. The first row shows results for k shuffles alone. The second row shows results for k shuffles followed by a binomial $(52, \frac{1}{2})$ cut.

With or without a cut, after five shuffles, the strategy gives two extra cards correct on average. After six shuffles, this goes down to about one, then it decreases by a factor of roughly two.

The strategy used is simple to describe and conjectured to be optimal for each k. If the deck has not been cut, guess the original top card as the first guess. As successive guesses are made and successive cards are revealed, check these cards off on a list of the deck in its original order. In general, such a list will have checked off cards and possible cards. Take a longest block of consecutive possible cards and guess the topmost card.

If the deck has been cut, the first card guessed is random and thereafter the above strategy is used starting with the first revealed card cycled to the top of the list. Both strategies are based on ideas of Michael McGrath.

5.2. *Analysis of shuffles with a random cut.* Michael McGrath has derived a simple closed form expression for the chance that a deck of n cards is in final

TABLE 5
Number of cards guessed correctly after k shuffles of 52 cards

k	1	2	3	4	5	6	7	8	9	10
No cut	31.17	19.69	12.92	8.80	6.56	5.51	5.01	4.76	4.65	4.60
Cut	29.45	19.09	12.69	8.70	6.50	5.46	4.97	4.73	4.63	4.57

arrangement π after k shuffles followed by a uniformly distributed cut. This chance is

$$\frac{a(\pi)\binom{2^k+n-r_1}{n}+b(\pi)\binom{2^k+n-r_2}{n}}{n2^{kn}},$$

where r_1 is the number of rising sequences in π when 1 is cut to top, r_2 is the number of rising sequences in π when 1 is cut to bottom, $a(\pi)$ is the distance from n to 1 counted forward and cyclically, $b(\pi)$ is $n - a(\pi)$. For example, with $n = 6$, $\pi = 5, 2, 3, 1, 6, 4$ has chance

$$\frac{5\binom{2^k+n-3}{n}+1\binom{2^k+n-4}{n}}{n2^{kn}}.$$

5.3. *Face up, face down shuffling.* As explained at the end of Section 3, similar analysis can be undertaken for some other groups. For the hyperoctahedral group, the shuffling operation has a simple interpretation: Cut the deck into two parts according to a binomial distribution. Turn the original top part face up and shuffle the two parts together by the Gilbert–Shannon–Reeds model. This basic shuffle is repeated k times. One is interested in the order of the cards becoming random as well as their face up, face down pattern becoming random. There are $2^n n!$ possibilities. The set of all such signed permutations forms a group usually denoted B_n. This is isomorphic to the group of symmetries of an n-dimensional cube.

Robert Beals has shown an elegant way in which analyses of such face up, face down shuffles reduces to the analysis of the Gilbert–Shannon–Reeds measure.

THEOREM 5. *Let $\overline{Q}$ denote the analog of the Gilbert–Shannon–Reeds measure on the group B_n, with $\overline{U}$ the associated uniform distribution. Let Q and U denote the GSR and uniform distribution on S_n. Then for $k = 1, 2, 3, \ldots$,*

$$\|\overline{Q}^k - \overline{U}\|_{B_n} = \|Q^{k-1} - U\|_{S_n}.$$

At the heart of Beal's argument is the following combinatorial fact: After two GSR shuffles of a deck of n cards, imagine straightening the cards out by removing the face up cards, keeping them in the same relative order, and placing them face down on top. The resulting random permutation has a GSR distribution.

A practical application arises in shuffling cards with oriented backs. There, one wants the arrangement of faces as well as the up-down pattern of the backs to be random. These orientations are significant for Tarot cards.

An analog of the algebraic results in Section 3 has been developed for B_n by Bergeron (1990) and Bergeron and Bergeron (1990, 1991).

Acknowledgments. Jim Pitman, Jim Reeds, David Aldous and Sophie Yancopoulos have spent hours discussing shuffling with us. Bob Beals, David Gay, Michael McGrath and Dan Rockmore all contributed results in Section 5. The geometric model of Section 3 was developed in conversations with Izzy Katznelson. Ira Gessel, Phil Hanlon and Richard Stanley have provided crucial links to combinatorical theory and commutative algebra. Laurie Beckett helped us simplify the proof of Theorem 1.

Note added in proof. Diaconis, McGrath and Pitman (1991) have found closed form expressions for a variety of events after an a shuffle. For example, the expected number of fixed points is $1 + 1/a + 1/a^2 + \cdots + 1/a^{n-1}$. The chance that π has n_1 fixed points, n_2 transpositions ... in its cycle decomposition is

$$\frac{1}{a^n}\prod_{i=1}^{n}\binom{f_i(a)+n_i-1}{n_i} \quad \text{with } f_i(a) = \sum_{d|i}\mu(d)a^{i/d}$$

where μ is the Möbius function of elementary number theory.

These results show that the approach to randomness has a rather delicate structure. The present paper shows that it takes $\frac{2}{3}\log_2 n$ shuffles to mix up n cards. The new results show that functions of π that only depend on cycles have approximately the right distribution after $c(n)$ steps, where $c(n) \nearrow \infty$ arbitrarily slowly. Preliminary computations indicate that functions which depend on large cycles (such as the length of the longest cycle) have the right distribution after 1 shuffle.

It is also worth noting that the algebraic results of Corollary 3 in Section 3 yield the eigenvalues of the basic Markov chain. The eigenvalues are $1, 1/2, 1/4, \ldots, 1/2^{n-1}$, with $1/2^i$ having multiplicity the number of permutations in S_n with $n - i$ cycles. This follows from Corollary 3: The expression for B_2^m only involves m through these powers of 2. From results in Hanlon (1990), the multiplicity of an eigenvalue equals the coefficient of the identity in the associated idempotent. Using the explicit expression for the idempotents completes the proof.

REFERENCES

ALDOUS, D. (1983). Random walk on finite groups and rapidly mixing Markov chains. *Séminaire de Probabilités XVII. Lecture Notes in Math.* **986** 243–297. Springer, New York.

ALDOUS, D. and DIACONIS, P. (1986). Shuffling cards and stopping times. *Amer. Math. Monthly* **93** 333–348.

BARR, M. (1968). Harrison homology, Hochschild homology and triples. *J. Algebra* **8** 314–323.

BERGERON, N. (1990). A decomposition of the descent algebra of the hyperoctahedral group II. *J. Algebra*. To appear.

BERGERON, F. and BERGERON, N. (1990). A decomposition of the descent algebra of the hyperoctahedral group I. *J. Algebra*. To appear.

BERGERON, F. and BERGERON, N. (1991). Orthogonal idempotents in the descent algebra of B_n. *J. Pure Appl. Algebra*. To appear.

DIACONIS, P. (1988). *Group Representations in Probability and Statistics*. IMS, Hayward, Calif.

DIACONIS, P., GRAHAM, R. and MORRISON, J. (1990). Asymptotic analysis of a random walk on a hypercube with many dimensions. *Random Structures and Algorithms* **1** 51–72.

DIACONIS, P., MCGRATH, M. and PITMAN, J. (1991). Cycles and descents of random permutations. Technical report, Dept. Statistics, Stanford Univ.

DIACONIS, P. and PITMAN, J. (1991). Analysis of top in at random shuffles. Technical report, Dept. Statistics, Stanford Univ.

DIACONIS, P. and ZABELL, S. (1982). Updating subjective probability. *J. Amer. Statist. Assoc.* **77** 822–830.

FELLER, W. (1971). *An Introduction to Probability and Its Applications* **2**, 2nd ed. Wiley, New York.

GARDNER, M. (1966). *Martin Gardner's New Mathematical Diversions from Scientific American*. Simon and Schuster, New York.

GARDNER, M. (1977). *Mathematical Magic Show*. Knopf, New York.

GARSIA, A. and REUTENAUR, C. (1989). A decomposition of Solomon's descent algebra. *Adv. in Math.* **77** 189–262.

GERSTENHABER, M. and SCHACK, S. (1987). A Hodge-type decomposition for commutative algebra cohomology. *J. Pure Appl. Algebra* **48** 229–247.

GESSEL, I. (1988). Unpublished lecture notes. Dept. Mathematics, Brandeis Univ.

GILBERT, E. (1955). Theory of shuffling. Technical memorandum, Bell Laboratories.

HANLON, P. (1990). The action of S_n on the components of the Hodge decompositions of Hochschild homology. *Michigan Math. J.* **37** 105–124.

HELLER, A. (1965). On stochastic processes derived from Markov chains. *Ann. Math. Statist.* **36** 1286–1291.

JORDAN, C. T. (1916). Long distance mind reading. *The Sphinx* **15** 57.

JORDAN, C. T. (1919). *Thirty Card Mysteries*. Published by the author, Pengrove, Calif.

LODAY, J. (1989). Operations sur l'homologie cyclique des algèbres commutatives. *Invent. Math.* **96** 205–230.

REEDS, J. (1981). Unpublished manuscript.

ROGERS, L. and PITMAN, J. (1981). Markov functions. *Ann. Probab.* **9** 573–582.

SOLOMON, L. (1976). A Mackey formula in the group ring of a Coxeter group. *J. Algebra* **41** 255–264.

STANLEY, R. (1977). Eulerian partitions of a unit hypercube. In *Higher Combinatorics* (M. Aigner, ed.). Reidel, Dordrecht.

STANLEY, R. (1986). *Enumerative Combinatorics*, Vol. 1. Wadsworth and Brooks/Cole Advanced Books & Software, Monterey, California.

TANNY, S. (1973). A probabilistic interpretation of the Eulerian numbers. *Duke Math. J.* **40** 717–722. [Correction (1974) **41** 689.]

THORP, E. O. (1973). Nonrandom shuffling with applications to the game of Faro. *J. Amer. Statist. Assoc.* **68** 812–847.

WILLIAMS, C. O. (1912). A card reading. *The Magician Monthly* **8** 67.

DEPARTMENT OF MATHEMATICS
COLUMBIA UNIVERSITY
NEW YORK, NEW YORK 10027

DEPARTMENT OF MATHEMATICS
HARVARD UNIVERSITY
CAMBRIDGE, MASSACHUSETTS 02138