MATHEMATICAL DEVELOPMENTS FROM THE ANALYSIS OF RIFFLE SHUFFLING

Persi Diaconis

1. *Introduction* The most common method of mixing cards is the ordinary riffle shuffle, in which a deck of n cards (often $n = 52$) is cut into two parts and the parts are riffled together. A sharp mathematical analysis for a natural model of riffle shuffling was carried out by Bayer and Diaconis (1992). This gives closed form expressions for the chance of any permutation and allows analytic approximation and exact numerical evaluation to show things like "seven shuffles are necessary and suffice to approximately randomize 52 cards". These results are carefully stated in Section 2A.

The shuffling work builds on earlier studies of Jordan (magic tricks), Borel (bridge), Gilbert, Shannon, Reeds (basic model) and D. Aldous (coupling). This background is described in Section 2B. The "seven shuffles" result is mildly dependent on the choice of metric and a number of alternative measures of randomness are discussed in Section 2C.

There is a mathematical reason that allows riffle shuffles to be analyzed so completely. The basic shuffling model falls squarely into Solomon's descent algebra (and indeed gives an independent development). This allows shuffling theorems to be translated into permutation enumeration results (e.g. how many permutations have a given number of descents and a given cycle structure). The eigenvalues of the Markov chain underlying shuffling were actually first determined in an investigation of Hochschild homology(Hanlon). There is an intimate connection with free Lie algebras and the Poincare-Birkoff-Witt Theorem (Bergeron-Bergeron-Garsia). The chance of a given cycle structure after riffle shuffling *equals* the chance that a random degree n polynomial has a given number of irreducible factors. This in turn is explained by considering the connection between shuffling and the action of the associated Lie type group $SL_n(F_q)$ on its Lie algebra (Fulman). Finally, shuffling gives a fairly direct interpretation of Schur symmetric functions (Stanley-Fulman). These results are described in section three.

The analyses above seem so rich and natural that they call out for generalization. A sweeping generalization of the theory was discovered by Bidigare, Hanlon and Rockmore. This involves random walk on the chambers of a hyperplane arrangement. The classical braid arrangement gives riffle shuffles but there are many other hyperplane arrangements where the chambers can be labeled by natural combinatorial objects and much (but not all) of the theory goes through. In an amazing synthesis, Ken Brown has shown that almost everything can be pushed through to random walks on idempotent semi-groups. This allows analysis of natural random walk on the chambers of spherical buildings. These results are described in Section 4.

The final section describes ten open problems.

Throughout I have tried to show the links with algebra. To be fair, many of the authors cited have no interest in the card shuffling implication of their work. This paper has improved from detailed comments of Ken Brown, Nantel Bergeron, Jason Fulman, Adriano Garsia and J.C. Uyemura-Reyes.

2A. Basic Shuffling.

The basic riffle shuffling model was introduced by Gilbert and Shannon (See Gilbert (1955)) and independently by Reeds (1981). It can be described as a probability distribution $Q(w)$ on the symmetric group S_n – the GSR Distribution. One description of Q is as follows: cut the deck into two piles according to the binomial distribution so the chance that pile one has j cards is $\binom{n}{j}/2^n$. Then, sequentially drop cards from the bottoms of the two piles according to the following rule: if at some stage pile one has "A" cards and pile two has "B" cards, drop the next card from pile one with probability $A/(A+B)$. This is continued until the two piles are exhausted and then the piles are pushed together. An equivalent description in terms of inverse riffle shuffles is due to Gilbert, Shannon and Reeds. An inverse shuffle begins by labeling each of n cards zero or one by a flip of a fair coin. Then, all the cards labeled zero are removed and placed on top keeping the cards in the same relative order. It is a simple exercise to show that the forward and backward descriptions are the same. From either description, given the cut, all ways of interleaving are equally likely, so the GSR shuffle is a maximum entropy model. The identity has probability $\frac{n+1}{2^n}$ while all other possible permutations have probability $1/2^n$.

Repeated shuffles are modeled by convolution:

$$Q * Q(w) = \sum_u Q(wu^{-1})Q(u)$$

Thus the chance of w after two shuffles is calculated as the chance of first choosing u and then choosing the permutation resulting in w. Similarly, $Q^{*k}(w) = \Sigma_u Q^{*(k-1)}(wu^{-1})Q(u)$. These ingredients complete the description of the GSR measure $Q^{*k}(w)$. Of course, shuffling is an example of random walk on a group and of a finite state-space Markov chain. See Saloff-Coste (2001, 2002), for an extensive overview with relevance to shuffling.

Repeated shuffling converges to the uniform distribution $U(w) = 1/n!$. The earliest works on Markov chains, due to Markov (1906) and Poincare (1912), used shuffling cards as an example. They gave results which allow us to conclude that

$$Q^{*k}(w) \to U(w) \quad \text{as} \quad k \to \infty.$$

It is natural to try to quantify this statement. The usual distance to stationarity is the total variation distance

$$\|Q^{*k} - U\| = \max_{A \subset s_n} |Q^{*k}(A) - U(A)| = \frac{1}{2} \sum_w |Q^{*k}(w) - U(w)|.$$

Consider the middle term. Its interpretation is this: let A be any subset of S_n (e.g. the set of all permutations where the ace of spades is in the top half). Calculate the chance of A after k shuffles (that is $Q^{*k}(A)$). Calculate the uniform measure of A (namely $|A|/n!$). Take the difference between these numbers and then take the A which makes this difference as large as possible. This is a very non-forgiving distance. If $\|Q^{*k} - U\| \leq \epsilon$ then shuffling is close to uniform for any set A.

These definitions translate the basic question "how many times should a deck of cards be shuffled to adequately mix it?" into a well posed math problem: given $\epsilon > 0$ how large should k be to have $\|Q^{*k} - U\| < \epsilon$? A historical review of progress on this problem is contained in the following section. Here we state the basic result.

Theorem 1. [Bayer-Diaconis]. When $n = 52$, the distance to uniformity is

k:	1	2	3	4	5	6	7	8	9	10
$\|Q^{*k} - U\|$	1.000	1.000	1.000	1.000	.924	.614	.334	.167	.085	.043

For general n, and $k = \frac{3}{2}\log_2 n + c$

$$\|Q^{*k} - U\| = 1 - 2\Phi\left(\frac{-2^{-c}}{4\sqrt{3}}\right) + 0\left(\frac{1}{n^{1/4}}\right) \quad \text{with} \quad \Phi(x)\frac{1}{\sqrt{2\pi}}\int_{\infty}^{x} e^{-t^2/2}dt.$$

Remarks The total variation distance is a number between zero and one. A graph of Q^{*k} vs k shows it stays close to its maximum until a bit before $\frac{3}{2}\log_2 n$ and then falls exponentially fast to zero. The analysis shows that for $k = \frac{3}{2}\log_2 n + c$ the distance goes to one doubly exponentially fast as $c \to -\infty$ and to zero exponentially fast as $c \to \infty$. These asymptotic results are borne out by the data for $n = 52$. The cutoff occurs at about seven shuffles. From six shuffles on, the variation distance falls by a factor of 2 for each extra shuffle. It is worth noting that the table is derived from exact results from Theorem 2 below not from an asymptotic approximation.

It is natural to wonder if this mathematics has much to do with real shuffles. People used to think that cards were suitably well mixed after three, four or five shuffles. Like many things people believe, this is simply not true. In Section 2B a classical card trick and some extensive analysis of bridge hands are used to prove this point.

Theorem 1 is a consequence of a more central result. To explain it, it is useful to have a geometric description of riffle shuffles. Picture n points dropped uniformly and independently into the unit interval. Label the ordered points, left to right, as $x_1, x_2, \ldots, x_n$. Now perform the baker's transform of $[0, 1]$ to itself. This takes $x \to 2x(\text{mod } 1)$. The points x_i are permuted inducing a permutation w. Note that there are a binomial number of the x_i in $[0, 1/2]$. The baker's map stretches these out to $[0, 1]$. The same holds for the points in $[\frac{1}{2}, 1]$. These two sets of points are interrleaved. It is not hard to see that the induced permutation has exactly the GSR distribution $Q(w)$.

This geometric description suggests a variant which will prove useful. For positive integer a, consider the a-shuffle which results from n random points under the map $x \to ax$ (mod 1). In shuffling language one may cut the deck into a-packets according to the multinomial distribution $\binom{n}{n_1 \ldots n_a}/a^n$ with $0 \le n_i \le n$, $\Sigma_{k=1}^{a} n_i = n$. The a packets are sequentially mixed, dropping a card from the ith packet with probability proportional to packet size. These equivalent definitions result in a probability $Q_a(w)$. In present notation $Q_2(w) = Q(w)$. From the geometric de-

scription, it is easy to see that an a-shuffle followed by a b-shuffle is the same as an ab shuffle thus $Q_a * Q_b = Q_{ab}$ and $Q_2^{*k} = Q_2 k$. It is enough to study only a-shuffles. The main result of Bayer-Diaconis can now be stated.

Theorem 2 For all positive integers n and a, $Q_a(w) = \dfrac{\binom{n+a-r}{n}}{a^n}$ with $r = r(w)$ the number of rising sequences in w.

To explain, consider a permutation w as an arrangement of a deck of cards, with w_i the label of the card at position i. Decompose w into disjoint rising sequences by finding card labeled 1, and then card labeled 2 if label 2 is below label 1. Continue until label k stopping if label $k+1$ is above one of $\{1, 2, \ldots, k\}$. Remove cards labeled $\{1, 2, \ldots, k\}$. This is the first rising sequence. Continue with the reduced deck, finding $\{k+1, \ldots, k+\ell\}$ a second rising sequence and so on. Thus, for $n = 9$, the permutation 716248359 has rising sequences 123, 45, 6, 789. Let $r = r(w)$ be the number of rising sequences obtained. Thus $1 \le r \le n$. Another description: $r(w) = d(w^{-1}) + 1$ with $d(w^{-1})$ the number of descents in w^{-1}. Descents will make a major appearance in Section 2C.

We will not give a proof of Theorem 2 here (see Bayer-Diaconis (1992) or the clear elementary treatment of Mann (1995)). It is straightforward from the geometric description using the "stars and bars" argument of elementary combinatorics. The hard part was discovering the result. We did this by looking at exact computer calculations for small decks (size 3, 4, 5) and noticing a pattern.

Theorem 2 gives yet another description of the GSR measure

$$Q(w) = \begin{cases} (n+1)/2^n & \text{If } w = id \\ 1/2^n & \text{If } w \text{ has 2 rising sequences} \\ 0 & \text{otherwise.} \end{cases}$$

Theorem 2 reduces the calculation of total variation to evaluating

$$\|Q^{*k} - u\| = \frac{1}{2} \sum_{j=1}^{n} k_n(j) \left| \frac{\binom{n+2^k-j}{n}}{2^{kn}} - \frac{1}{n!} \right|$$

with $k_n(j)$ the number of permutations with j rising sequences. At this point another surprise occurred. The $k_n(j)$ are very well studied as the coefficients of Eulerian polynomials (Stanley, 1997). This allowed careful asymptotic analysis and led to Theorem 1.

This concludes our overview of the basic shuffling story. We turn to a bit of history and then some more mathematical consequences.

2B. History and Practical Consequences.

The earliest treatments of Markov chains treat card shuffling as a leading example (Markov (1906), Poincare (1912), Doob (1954)). These treatments show that shuffling cards *eventually* results in a well mixed deck. It is very hard to guess

how many shuffles are needed. When $n = 52$, $n! \doteq 8 \times 10^{67}$. For the other popular method of shuffling (overhand) Pemantle (1984) shows order $n^2 \log n$ shuffles suffice. This is more than 2500 when $n = 52$.

Emil Borel in Borel and Cheron (1940) began a quantitative investigation by studying how long individual cards and pairs of cards take to randomize. This allowed him to conclude that at least six or seven shuffles are needed. Similar conclusions are drawn by Keller (1995).

Independently, magicians had discovered that rising sequences allowed good card tricks to be performed if cards are not well shuffled. Details and references appear in Bayer-Diaconis (1992). Bridge players went from hand shuffling to computer generated shuffling in their tournaments. A comparison of before and after suit distribution shows that the standard four or five riffle shuffles followed by a cut are grossly inadequate Berger (1973). Thorpe (1972) is an early survey detailing ways of taking advantage of poor shuffling in casino games.

Modern work on the mathematics of riffle shuffling begins with work of Gilbert (1955). He reports joint work with Shannon on the GSR model. They proved some combinatorial properties of GSR shuffles and suggested $\log_2 n$ would be enough. The model was independently discovered by Reeds (1981) who made extensive computer studies. Aldous (1983) gave a coupling argument which proves that $\frac{3}{2} \log_2 n$ shuffles are sufficient for n large. Aldous and Diaconis (1986) carefully prove that

$$\|Q^{*k} - U\| \leq 1 - \prod_{i=1}^{n-1} \Big(1 - \frac{i}{2^k}\Big).$$

This bound becomes less than $\frac{1}{2}$ for $k = 11$ when $n = 52$.

An empirical study of the GSR model compared to actual shuffles appears in Diaconis (1988). This concludes that the model is a good fit. Of course, much depends on the shuffler – casino dealers (along with the present author) can shuffle close to perfectly and eight perfect shuffles recycle the deck! See Diaconis-Graham-Kantor (1983) or Morris (1990) for more of this. There is much further work to do in developing tractable models with a few parameters which allow individual tuning. Because of its maximum entropy property the GSR model offers a provable lower bound to any less uniform distribution.

As a final practical note, Diaconis-Holmes (2000) analyze a class of mechanical 'shelf-shufflers' used in casino games. In these, a deck of n cards is distributed randomly onto a shelves. At each stage, cards are placed at random above or below previously placed cards on a shelf. At the end, the packets are output in random order (it turns out not to matter). The shuffle is *not* repeated. It turned out that the theory developed for type B (hyperoctahedral group) gave a complete analysis.

2C. Other Measures of Randomness

The results of Theorem 2 allow computation in various alternatives to the total variation metric. Aldous and Diaconis (1983) derive results for separation distance

$$s(k) = \max_w 1 - \frac{Q^{*k}(w)}{U(w)} = 1 - \frac{n!\binom{2^k}{n}}{2^{nk}} = 1 - \prod_{i=1}^{n-1} \Big(1 - \frac{i}{2^k}\Big).$$

As discussed above this needs $k = 11$ to make it small when $n = 52$. Su (1995), Trefethan-Trefethan (2002) and Stark et al. (2000) derive results for entropy distance that suggest $k = 5$ or 6 shuffles suffice when $n = 52$. The theorem of Stark et al. (2000) shows that the entropy distance decreases by a constant factor up to $\log_2 n$ shuffles when it goes to zero exponentially. A graph of the distance versus entropy for small values of n seems to show a discontinuous derivative at $\log_2(n)$. If true, this would be a new kind of phase transition. Lovasz and Winkler (1995) use Theorem 2 to show that a very different distance, the expectation of the fastest strong stationary time will be small after $k = 11$.

All of the above are global measures of uniformity. In explaining the convergence results to a popular audience, the following notion seemed useful. Consider playing the following game. A deck of cards is on the table. Guess at the top card. This card is then shown and discarded. Then guess at the next card (which is then shown and discarded) and so on. If the deck is perfectly mixed, the chance that the first guess is correct is $1/n$, the chance the second guess is correct is $1/(n-1)$, etc. Thus $1/n + 1/(n-1) + \ldots + 1$ correct guesses are expected. When $n = 52$ this is about 4.5. Suppose instead that k riffle shuffles have been carried out. A conjectured optimal strategy for guessing was derived by McGrath (see Bayer-Diaconis (1992)). Using the strategy yields about 5.01 correct guesses after seven shuffles with 4.97 correct following seven shuffles and a cut. In related work, Ciucu (1998) studies the optimal guessing strategy following k-riffle shuffles when *no* feedback is given. He proves that for $2n$ cards if $k \geq 2\log_2(2n) + 1$, the best strategy is to guess card 1 for the first half of the deck and card $2n$ for the second half. For $k < 2\log(2n)$, there are better strategies. In particular, after one shuffle he shows that guessing $1, 2, 2, 3, 3, 4, 4, \ldots$ in order gives $\sqrt{8n/\pi}$ correct guesses asymptotically. His analysis rests on an explicit diagonalization of the Markov chain which tracks the position of the card labeled 1. This is closely related to work in Section 4B below.

The above study suggested looking at classical permutation enumeration questions (e.g. number of fixed points or cycles) after an a-shuffle. This turned out to be surprisingly neat. For example, the expected number of fixed points is

$$E_a(Fp) = 1 + \frac{1}{a} + \frac{1}{a^2} + \ldots + \frac{1}{a^{n-1}}.$$

For cycles, the full story was derived by Diaconis-McGrath-Pitman (1995). Let $Q_a(n_1, n_2, \ldots n_n)$ be the chance that an a-shuffle results in a permutation with n_i i-cycles. They proved

$$Q_a(n_1 \ldots n_n) = \frac{1}{a^n} \prod_{i=1}^{n} \binom{n_i + f_i(a)}{n_i} \quad \text{with} \quad f_i(a) = \frac{1}{i} \sum_{d|i} \mu(d) a^{i/d} \tag{2.1}$$

The proof uses a remarkable bijection of Gessel and indeed gives a self-contained proof of Gessel's results – see Gessel-Reutenauer (1993) for Gessel's version with extensive application to enumerating permutations by descents and cycle structure. The formula 2.1 and some analysis show that features of a permutation that only depend on cycle structure become random before $\frac{3}{2}\log_2 n$-shuffles; the length of the longest cycle is close to its uniform distribution after one shuffle.

In a different direction, discussed further in Section 3B, Fulman (2002) has shown that the length of the longest increasing subsequence has its correct limiting

distribution after $\frac{5}{6}\log_2 n$ shuffles. These results also imply that the patience sorting solitaire game described by Aldous-Diaconis (1995) will then behave as if the deck was random.

Uyemura-Reyes (2002) has studied the number of riffle shuffles required to randomize just a few cards e.g. the original top card. He derives bounds using coupling and remarkable formulas for how the eigen-values of the GSR shuffles split by representations. His results generalize earlier profound work of Bergeron, Bergeron, Garsia (1989), and Hanlon (1990). They are discussed further in 4B below.

All of this shows that "seven shuffles suffice" is just a rough guide. From Theorem 1, it is where the cutoff happens.

To finish off this part of the shuffling story we note that the analysis has been broadened to show that the age old custom of following shuffling by a random cut does not help appreciably in convergence. This is illustrated in Bayer-Diaconis (1992) and much more sharply in Fulman (2000B). This last paper connects shuffling with cuts to cyclic descent theory.

3. Some Mathematical Connections

A. Descent Theory

A permutation w has a descent at i if $w_{i+1} < w_i$. The set of all such i makes up the descent set $D(w) \subseteq \{1, 2, \ldots, n-1\} = [n-1]$. Descents record the up down pattern in permutations and are a natural object of combinatorial study. Stanley [1972, 1986] lays out the classical theory and Buhler et al. (1994) make a fascinating connection to the mathematics of juggling. Stadler (1997) develops links between descents, shuffling and juggling for permutations of multisets.

Let $S \subseteq [n-1]$ and let $a_S = \sum_{w: D(w)=S} w$. Louis Solomon (1976) observed that as elements of the group algebra $Q[S_n]$, the a_S are the basis for a subalgebra now called Solomon's descent algebra. In particular $a_S a_T = \Sigma_u c_{ST}^u a_u$ for $c_{ST}^u \in \mathbf{Z}$. Solomon's motivation was to give a group theoretic interpretation of Mackey's induction theorem. He did this in a unified way for classical Weyl groups. The development he started now has a life of its own.

The connection to shuffling cards comes through the following observations. The set of permutations with a single descent at position i (along with the identity) are exactly the permutations realized by removing an i element subset of $1, 2, \ldots, n$ and placing them to the left (keeping all else in its same relative order). This is exactly the inverse riffle shuffles consonant with i cards cut. Summing in i, let $A_1 = \Sigma_{i=1}^{n-1} a_i$ this is the sum of all permutations with a single descent. Excepting the identity, it is also the result of an arbitrary inverse riffle. If Q is the Gilbert-Shannon-Reeds measure then, as an element in $Q[S_n]$,

$$\sum_w Q(w^{-1})w = \frac{n+1}{2^n}\ id + \frac{1}{2^n} A_1.$$

Thus the neat convolution properties of the GSR measure show that if A_i is the sum of permutations with exactly i descents (so $A_0 = id$), then $A_0, A_1, \ldots, A_{n-1}$

are a basis for a commutative subalgebra of the descent algebra. In particular, $A_i A_j = A_j A_i = \Sigma c_{ij}^k A_k$. This commutative subalgebra of the descent algebra appears in Bayer-Diaconis (1992). As explained there, close relatives had been discovered by Gerstenhaber-Schack [1987] in their development of Hochschild Homology and by Loday [1988], Hanlon [1990] in their development of cyclic homology. The idempotents of this algebra act naturally on a complex constructed from the usual bar resolution and, for commutative algebras, commute with the boundary maps. Hence their kernel and image offer a natural Hodge-type splitting of the associated homology.

It would take us too far afield to explain the connections between the descent algebra, the free-Lie algebra, and Philip Hall's commutator calculus. Fortunately, this has been splendidly carried out by Garsia (1990) and Garsia-Reutenauer (1989) as summarized by Reutenauer [1993]. This book contains a central chapter on shuffle algebras. It omits most of the topics discussed in the present review! A number of other appearances of shuffling are in the series of papers by Nantel Bergeron (with several sets of coauthors) listed in the bibliography. These extend previous results to more general Coxeter groups, include applications to Vassiliev invariants and much else.

B. Connections with Symmetric Function Theory

The theory of symmetric functions as developed by Stanley (1972, 1999) and Macdonald (1985) has had a great unifying effect on combinatorics. Many seemingly isolated facts about balls in boxes, permutations and partitions are nowadays seen as formulae for change of basis. Schurs symmetric functions are at the heart of this theory. A charming discovery of Stanley (2001) developed by Fulman (2002) shows how Schur functions arise in a natural way from riffle shuffling. Let $\theta_1, \theta_2, \ldots$ be non-negative numbers that sum to one. Drop n balls into a set of boxes with θ_i the chance of a ball dropping into box i. Suppose the box counts are $N_1, N_2, \ldots$ with $N_1 + N_2 + \ldots = n$. Take a deck of n cards; cut off the top N_1, cards, then the next N_2 cards (forming a separate pile), etc. of course, many of the piles may be empty. Riffle shuffle these piles together as in Section 2a. This results in a final permutation w. Apply the Schensted map to w to get a pair of standard Young-tableaux of the same shape λ.

Proposition The probability that the above procedure results in the partition λ is the Schur function s_λ times the dimension f_λ of the associated representation of the symmetric group:

$$s_\lambda(\theta_1, \theta_2, \ldots) f_\lambda.$$

Stanley's proof of this proposition uses quasi-symmetric functions, an emerging tool in algebraic combinatorics. Fulman's proof of the proposition uses only classical facts from symmetric function theory. Both authors develop corollaries and variations. One striking application to shuffling due to Fulman shows that the distribution of features of a permutation dependent on the shape of the associated Young-tableaux- e.g. the longest increasing subsequence – have the correct limiting distribution after $\frac{5}{6} \log_2 n$ shuffles. Stanley (1999) (2002) is a good place to start reading about quasi-symmetric functions. Aguiar and Sottile (2002), Billera, Hsiao

and Van Willigenburg (2001) and Garsia, Wallach (2002) are relevant, significant studies. All have shuffles as part of their combinatorial essence.

3C Work of Fulman Some profound connections between shuffling and the enumerative theory of finite groups of Lie type have been developed by Jason Fulman. Some of this has already made an appearance above in Sections 2B and 3B. This section describes some further developments. Yet others appear in the rich collection of papers listed in the bibliography.

One striking result of Fulman explains a mystery. A main result in Diaconis-McGrath-Pitman (1995) is a closed formula for the cycle structure of a permutation after an a-shuffle (see (2.1) in Section 2C). It was also observed that this formula answers a second question: pick a random monic degree n polynomial $x^n + a_{n-1}x^{n-1} + \ldots + a_0$ with coefficients in $\mathbf{F}_q$ by choosing $a_0, a_1, \ldots, a_{n-1}$ from the uniform distribution. Factor this polynomial into irreducible factors and suppose there are n_i irreducibles of degree i. The chance of a given $n_1, n_2, \ldots, n_n$ occurring is given by (2.1) with $a = q$. This was proved by observing that two formulae agreed – that is, without understanding. Fulman [1998] found a conceptual explanation and an extension to other groups and shuffling schemes.

Fulman's explanation begins with a simply connected, semi-simple group G defined over $\mathbf{F}_q$. Let $\mathcal{G}$ be the Lie algebra. Consider the orbits of semi-simple elements of $\mathcal{G}$ under the adjoint action of $\mathcal{G}$. For example, for groups of type A, $G = SL_n(\mathbf{F}_q), \mathcal{G} = s\ell(n, q)$ and semi-simple elements correspond to monic degree n polynomials with coefficient of x^{n-1} vanishing. For types A and B, Fulman shows that there is a natural map Φ from the semi-simple orbits to the conjugacy classes of the Weyl group W such that a uniformly chosen orbit maps to the measure induced by a-shuffling with $a = q$. Thus a randomly chosen polynomial maps to an a-shuffle and the factors map to cycles. For shuffles of type B, the correspondence is with symmetric polynomials $f(z) = f(-z)$

In algebraic group theory there is an analog of the map Φ which carries semi-simple conjugacy classes of the group G to conjugacy classes of the Weyl group. Picking a semi-simple class uniformly induces a probability distribution on conjugacy classes. Fulman [1997] managed to find a card shuffling interpretation of this map as well and give an enumerative theory that works for all split semi-simple groups. His work uses results of Cellini and Carter's work on the Brauer complex. Indeed, Carter (2002) has recently extended Fulman's work to more general groups.

We give the card shuffling version of Fulman's work for type A. Define an F-shuffle of a deck of $2n$ cards as follows: choose an even number j, between 1 and $2n$ with probability $\binom{2n}{2j}/2^{2n-1}$. Remove the top j cards of the deck. Remove the bottom j cards of the deck and place them on top of the original top j cards to form a packet of size $2j$. Shuffle this packet with the remaining $2n - 2j$ cards. Fulman derives remarkable closed form generating functions for the cycles of a permutation after an F-shuffle. He also shows that F-shuffles convolve nicely and, for special deck sizes, gives an alternate description in terms of a riffle followed by a cut.

The analogous developments for type B yield closed formulae for the cycles of randomly chosen unimodal permutations. These arise in dynamical systems and in social choice theory.

One further aspect of Fulman's work deserves special mention (and follow-up!). The shuffling work in Diaconis-McGrath-Pitman (1995) leans on a remarkable bijection of Gessel between multisets of primitive necklaces and permutations with cycle structure equal to that of the necklace. Fulman shows that by refining the correspondence Φ described above to a map to the Weyl group (instead of just to conjugacy classes) one recovers Gessel's bijection in a group theoretically natural way.

3C. Work of Lalley

Steve Lalley has written a series of papers studying extensions of the basic Gilbert-Shannon-Reeds model to less uniform methods of riffle shuffling. Even changing the method of cutting the deck in two from a fair binomial distribution to a skewed binomial distribution with parameter $p < \frac{1}{2}$ destroys a basic symmetry. For this case, Lalley [2000] conjectures that there is a sharp threshold for the mixing time at $C_p \log n$ for $C_p = (3 + \theta_p)/\log(1/p^2 + q^2)$ with θ_p the unique solution of $p^\theta + q^\theta = (p^2 + q^2)^2$. Observe that $C_{\frac{1}{2}} = \frac{3}{2}\log_2 n$ in agreement with Theorem 1. Lalley [2000] and Fulman (1998) give upper and lower bounds of this form for the mixing time but sharp results are conjectural.

Lalley [1996], [1999] expands the basic interlacing mechanism underlying the GSR shuffle. To explain, recall the dynamical systems description of GSR shuffles as the permutation induced by n uniform points in $[0, 1]$ under the baker's transformation $x \longmapsto 2x \bmod(1)$. This results in all interleaving being equally likely. It is natural to consider more general maps $f : [0, 1] \to [0, 1]$ which preserve Lebesgue measure. Lalley works with piecewise C^2 maps which are piecewise monotone increasing. He shows that several interpretable shuffles can be so described. For example, the biased cut shuffles described above or shuffles where the left card is dropped with probability $uA/(uA + wB)$ when packets are of size A, B, here u, w are fixed parameters. When $u = w = \frac{1}{2}$ this becomes the original GSR shuffle.

The main result of Lalley [1996] shows that when n is large, for fixed i, the number N_i of cycles of length i after an f-shuffle are approximately independent geometric random variables with $P(N_i = k) = (1 - w)w^k$ the parameter w depends on i and on the map f in a simple way. Further, the N_i are approximately independent. The main result of Lalley [1999] gives a lower bound for the number of f-shuffles required to mix N cards; at least $h^{-1} \log N$ shuffles are needed where h is the 'fiber entropy' associated to f. The proofs are a marvelous mix of ergodic-theoretic symbolic dynamics and combinatorics.

One interesting aspect of these f-shuffles is that, aside from a-shuffles, the successive permutations chosen for repeated convolution are not independent. They form a stationary sequence. This is not necessarily bad; perhaps real shufflers remember a few steps back – if a particularly lumpy shuffle was just made the next shuffle might be neater. See also Dubrow-Fill (1995). There is much to follow up from Lalley's work. Perhaps the leading problem is to prove any kind of upper bound for f-shuffles or better, to determine where cutoffs appear.

3D. Early Shuffling

The basic combinatorial shuffling of two sequences, one with m letters $x_1, \ldots, x_m$

and one with n letters $y_1, \ldots, y_n$, into the formal sum of sequences of $n+m$ letters in all orders that preserve the order of the x's and the order of the y's (thus $\binom{n+m}{m}$ terms) appears in other areas of algebra.

Perhaps earliest is the classical wedge product of two alternating forms. If V is a vector space and $f : V^m \to \mathcal{R}$, $g : V^n \to \mathcal{R}$ are alternating multilinear functions, then $f \wedge g : V^{n+m} \to \mathcal{R}$ may be constructed as the function

$$f \wedge g(x_1, \ldots, x_{n+m}) = \sum_{\sigma} sgn(\sigma) f(x_{\sigma_1}, \ldots, x_{\sigma_m}) g(x_{\sigma_{m+1}}, \ldots, x_{\sigma_{n+m}})$$

where the sum is over all shuffles. A splendid account of this classical subject appears in Cartan (1967, pg. 179-188). The shuffling construction guarantees that $f \wedge g$ is alternating, that $f \wedge g = (-1)^{mn} g \wedge f$ and that the wedge product is associative. Cartan's proof of this last statement results from the following fact: with three packets of cards of sizes ℓ, m, n, shuffling ℓ into m and then the n into this joint packet results in the same distribution as shuffling in any of the other orders, or indeed shuffling the 3 packets together simultaneously as in the 3-shuffles described in Section 2d. More general shuffles appear when studying flag manifolds. A flag is an increasing sequence of subspaces. If the successive dimensions are $n_1, n_1+n_2, \ldots$ then shuffles based on cutting off packets of size $n_1, n_2, \ldots$ appear. In particular, such shuffles index a basis for the homology of the associated flag variety. See Fulton (1997) or Shahshahani (2002) for textbook descriptions.

Eilenberg-MacLane (see MacLane 1950) used the shuffle construction as a basic building block for constructing a chain complex giving an appropriate cohomology theory for Abelian groups. They get $H^2(\pi, G)$ as the group of Abelian extensions of G by π.

Shuffles appear frequently in other basic constructions in algebraic topology. For example, if X is a space with an associative, commutative product, Milgram (1967) defined a product on the classifying space $B(X)$ using shuffles. This work was systematized by Steenrod (1967) and further by MacLane (1970). Shuffles appear in the Eilenberg-Zilber Theorem and in explicit proofs of the Künneth formula giving a chain equivalence between a chain complex for the product of two spaces and the tensor product of the two chain complexes. See Hatcher (2002, pg. 278) for details and Dupont (2001, pg. 29) for a charming appearance in the world of scissors congruences! The essence of much of this is that the shuffling map gives a natural triangulation of the product of two simplices.

¿From a modern view, many of these appearances of shuffling occur because of the many natural Hopf algebras in mathematics. See Schneider-Sternberg (1993) for references and pointers to Rees' shuffle algebras and Chen's iterated integrals. Perhaps even more basic, the permutatedron is the convex hull of all permutations of the vector $(1, 2, 3, \ldots, n)$ in $\mathbf{R}^n$. It is a convex polytope with vertices indexed by permutations. It may be seen that the edges and faces of various dimensions are indexed by shuffles. See Billera and Sarangarajan (1996) for a clear statement and proof. It would be marvelous if some of what we know about shuffling illuminates these applications or vice versa.

4. Some Generalizations

There are a bewildering variety of extensions of riffle shuffling where much of the successful analysis goes through. It is easiest to lead into this by considering inverse riffle shuffles where a subset is selected at random and moved to the top. A natural generalization is to partition $[n]$ into ordered blocks $(B_1, B_2, \ldots, B_k)$. Then remove all cards with labels in block one and move these to the top (keeping the cards within a block in their original relative order). Next cards with labels in B_2 are removed and put directly below those in B_1, and so on with cards having labels in block k finishing at the bottom. Let $\mathcal{B}$ be the space of all ordered blocks of any shape if a weight $w(B), B \in \mathcal{B}$ is specified with $w(B) \geq 0$ $\Sigma w(B) = 1$, then a random walk can proceed.

Inverse riffle shuffles and the GSR model proceed from the uniform distribution on the set of 2^n partitions into two blocks. A widely studied special case puts weights $w_1, w_2, \ldots, w_n$ on each card and then removes card i to top. This arises as a method of rearranging files so that frequently called for items are near the top. See Fill [1996] for an extensive survey. Curiously, the special case with $w_i = 1/n$ for all i is central in Wallach (1986) and Garsia-Wallach (2002).

As will emerge, there is a relatively complete theory for this class of walks – a description of stationary distribution, reasonable rates of convergence and a complete description of the associated eigenvalues. This will follow from the following sweeping generalization.

A. Hyperplane Walks

Bidigare-Hanlon-Rockmore (1999) introduced a class of walks on chambers of a hyperplane arrangement which includes the walks above as a special case. Their works was completed in various ways by Bidigare (1997), Brown [2000, 2001], Brown-Diaconis [1998]. Billera-Brown-Diaconis (1999) offer an introduction.

The story begins with a set $\mathcal{A}$ of affine hyperplanes in $\mathbf{R}^d$. This cuts $\mathbf{R}^d$ into regions called chambers. These chambers are polyhedra with sides called the faces of the arrangement. For example, three lines in the plane in general position yield 7 chambers (2-dimension), 9 one dimensional faces and three zero-dimensional faces (the three points of intersection). Given a face F and a chamber C, the projection of F on C, written FC, is the unique chamber with F as a face and closest to C. Here closeness if measured by the number of hyperplanes in $\mathcal{A}$ one must cross in moving from C to FC.

Let $w_F \geq 0$ $\Sigma w_F = 1$ be weights on the faces of the arrangement $\mathcal{A}$. Define a random walk on the set of chambers by moving from C to FC when F is chosen with probability w_F. The theory depends on the lattice $\mathcal{L}$ of all possible intersections of elements in $\mathcal{A}$. Here are the main theorems of Bidigare-Hanlon-Rockmore [1999], Brown-Diaconis [1998].

Theorem 1 Let $\mathcal{A}$ be a hyperplane arrangement in $\mathbf{R}^d$. Let $\mathcal{L}$ be the intersection lattice of $\mathcal{A}$ and w_F a probability measure on the faces. Then, the transition matrix of the Markov chain is diagonalizable. For each $W \in \mathcal{L}$ there is an eigenvalue

$$\lambda_W = \sum_{F \leq W} w_F$$

with multiplicity

$$m_W = |\mu(W, V)| = (-1)^{\dim(W,V)}\mu(W, V)$$

where μ is the Möbius function of $\mathcal{L}$.

Theorem 2

(a) The Markov chain of Theorem 1 has a unique stationary distribution π if and only if for each $H \in \mathcal{A}$ there is a face F not in H with $w_F > 0$.

(b) The stationary distribution in (a) can be described by sampling faces *without replacement* from w_F to get an ordering $F_1, F_2, \ldots$. Then, for any chamber C_0, the product $F_1F_2F_3 \ldots C_0$ is a chamber distributed from π.

(c) For π as in (a), (b), and starting chamber C_0

$$\|K^\ell_{C_0} - \pi\| \leq \sum_{H\in\mathcal{A}} \lambda^\ell_H$$

To complete this section, let us show how these hyperplane walks extend riffle shuffles. The *braid arrangement* $\mathbf{A}_d$ consists of hyperplanes $H_{ij} = \{x \in \mathbf{R}^d : x_i = x_j\}$. All points within the same chamber have the same relative order so the chambers may be labeled with permutations. The faces are points in $\mathbf{R}^d$ which lie on some of the H_{ij} and on various sides of the rest. These may be labeled by block ordered partitions $(B_1, B_2, \ldots, B_k)$ of $[n]$. Finally, the action FC of a block ordered partition on the permutation corresponding to C results from removing cards from the first block and moving to the top, etc., as described in the introduction to this section.

The present description does not do justice to the wealth of examples of hyperplane arrangements where the chambers have natural names and the walk has a natural interpretation. We can only hope that the reader will consult the references above.

B. Some Representation Theory

I want to describe work of Bergeron-Bergeron-Garsia (1989), Hanlon [1990] and Uyemura-Reyes (2002) which shows a deep interplay between the shuffling schemes of Section A and the representation theory of the symmetric group. To keep things manageable, consider random walks on the braid arrangement driven by invariant face weights: $w(F) = w(\sigma F)$. This includes (uniform) random to top and inverse riffle shuffles as special cases. Let $Q(\sigma) = \Sigma_{Fid=\sigma} w(F)$. These walks may be described via repeated convolution by the probability measure Q.

It is natural to ask how the eigenvalues of the walk split up by representation. Recall that the irreducible representations of S_n are indexed by partitions ν of n. If $\rho_\nu(\sigma)$ is the associated matrix representation, we are asking about the eigenvalues of the matrix $\widehat{Q}(\nu) = \Sigma_\sigma Q(\sigma)\rho_\nu(\sigma)$. By general theory (Diaconis (1988, Chapter 3E)) these are a subset of the eigenvalues from Theorem 1 in Section 4A Above. For the

braid arrangement, the eigenvalues are indexed by block ordered partitions. However, because of the symmetry $w(F) = w(\sigma F)$, the eigenvalues only depend on the underlying number partition. Thus for each pair of partitions (μ, ν) we may ask how many times the eigenvalue λ_μ occurs in the matrix $\widehat{Q}(\rho_\nu)$. To describe the answer we need both the usual irreducible characters χ_ν of S_n and the Lie characters ψ_μ (Reutenauer (1993, Chapter 8)). These Lie characters may be described by taking a permutation of cycle type μ in S_n. Its centralizer is a product of Wreath products $S_k wr C_j$. Take a ξ primitive jth root of 1, consider the one dimensional character of C_j^k which takes $x_j, \ldots, x_n$ to $\xi^{x_1+\cdots+x_k}$. This induces a one dimensional character of the Wreath Product. Taking a product of these 1-dimensional characters over all factors in the centralizer and then inducing up from the centralizer to S_n gives ψ_μ. The main theorem below was proved by Hanlon [1990] for the case of GSR shuffles. Richard Stanley (personal communication) conjectured the general result which was proved by Uyemura-Reyes (2002).

Theorem 3

For an S_n invariant hyperplane walk on the braid arrangement the multiplicity of the eigenvalue λ_μ of Theorem 1 in the ν^{th} irreducible representation of S_n equals

$$\langle \chi_\nu, \psi_\mu \rangle.$$

Remarks

(a) Lie characters have been extensively investigated when $\mu = (n)$, see Stembridge (1989), where an explicit decomposition formula is given. For general partitions μ, much less is known. Theorem three shows that the S_n invariant shuffles are equivalent objects in the group algebra. Any such shuffle is a linear combination of what may be called μ shuffles as described in the introduction to this section. As shown in Diaconis-Fill-Pitman [1992, Sec. 5], these μ shuffles form a basis for the descent algebra.

(b) Uyemura-Reyes (2002) shows how the numbers described above allow bounds on how many shuffles of a given type are required to randomize a subset of cards, e.g. the original top card or top 13 cards. Here is one example. If $k = \log_2(n/c)$, after k inverse GSR shuffles, let Q_k be the probability distribution of the position of the original top card. Then Q_k is close to uniform if c is small:

$$\|Q_k - u\| \leq 1 - (1 - 2^{-k})^n$$

(c) These connections to representation theory are crucially used in Fulman (2000B) to get nice formulae for the cycle structure of shuffles followed by a cut.

C. Brown's Semigroup Walks

Ken Brown [2000, 2001] has given a marvelous extension of the hyperplane walks which leads to interesting special cases *and* a conceptual explanation of why the

eigenvalues of these non-symmetric Markov chains are non-negative real numbers. The brief treatment given here is a shuffling together of two of Brown's papers and the reader is strongly encouraged to read the originals.

Let S be a finite semigroup satisfying $x^2 = x$ for all $x \in S$. A random walk is driven by a probability distribution $w(x), x \in S$. At each stage, one picks x from $w(x)$ and multiplies on the left. Thus the transition matrix is

$$K(s,t) = \sum_{x \cdot s = t} w(x)$$

In all the examples, the state space of the walk is restricted to a left ideal I in S.

Example 1. Hyperplane Walks. Let S be the set of faces of a hyperplane arrangement with I the set of chambers under the product of Section 4A. This product is idempotent and the results of Section 4A will be seen as special cases of the main theorem below.

Example 2. q analogs Let $MAT(n, \ell, q)$ be the set of $n \times \ell$ matrices of rank ℓ with coefficients in $\mathbf{F}_q$. Let $S = \cup_{\ell=1}^{n} MAT(n, \ell, q)$ and $I = GL_n(q) = MAT(n, n, q)$. Define a product on S as follows: If s has columns $(s_1, \ldots, s_\ell)$ and t has columns $(t_1, \ldots, t_m)$ form $s \cdot t$ by appending the columns of t to the columns of s in order $t_1, t_2, \ldots$ deleting a t_i if it is linearly dependent on the columns already there. This is an idempotent associative product and $GL_n(q)$ is an ideal.

The "$q = 1$ case" consists of ordered strings from $1, 2, \ldots n$, without repeated values and the ideal becomes the symmetric group S_n. Thus if $s = (3, 5)$ and $t = (23145)$ $st = 35214$ and we see the move to the front chain.

Example 3. The free idempotent semigroup F_n on $1, 2, \ldots n$, may be described as the equivalence class of finite strings under the equivalence relations $w^2 = w$ for all subwords. For example, when $n = 2$, we get the six strings

$$S = \{1, 2, 12, 21, 121, 212\}$$

Brown (following Green and Reees (1952)) shows that F_3 has order 159 and F_n has order $\sum_{i=1}^{n} \binom{n}{i} \prod_{j=1}^{i} (i - j + 1)^{2^j}$.

Let I be the ideal of all words having each of $\{1, 2, \ldots, n\}$ appearing at least once (for $n = 2, I = \{12, 21, 121, 212\}$). Any probability measure on S induces a Markov chain on I by left multiplication.

Return now to the general case of an idempotent semigroup S. Brown introduces a support map supp : $S \to L$ with L an explicitly constructed semilattice. The support map is a subjection satisfying supp $(xy) =$ supp $x \vee$ supp y and supp $x \geq$ supp y if and only if $x = xyx$. The set L indexes the eigenvalues of the walk. For hyperplane walks, L is the intersection lattice. For matrices, L is the subspace spanned by the columns. For the free idempotent semigroup L is the collection of subsets of $\{1, 2, \ldots, n\}$ under union. The natural ideal I is the two sided ideal. $\{x : \text{supp } x = \widehat{1}\}$. This specializes to the ideals given in the three examples above.

Brown gives a version of theorems one and two of Section 4A: For each $X \in L$ there is an eigenvalue $\lambda_X = \sum_{\text{supp } x \leq X} w(x)$ with a neat way of computing multiplicities. If the product of x with $w_x \neq 0$ is in I then there is a unique

stationary distribution π which may be described as the distribution of the random element $x_1 x_2 \dots c_0$ with $x_1, x_2, \dots$ sampled *without* replacement from $w(x)$. Finally, for any starting state $C_0 \in I$,

$$\|K_{C_0} - \pi\| \leq \sum_H \lambda_H^\ell$$

where H ranges over the maximal elements of L.

The key to the analysis is a surprising, complete character theory. (Most semigroups do not have a reasonable character theory.) Brown shows that all representations of S are one dimensional and that the representations are indexed by L; the 'Fourier transform' of the random walk now yields the eigenvalues.

One aspect of Theorem 1 that needn't go through: the Markov chain needn't be diagonalizable. To help the reader navigate, Brown first worked in idempotent semigroups satisfying the additional identity $xyx = xy$. These are called left regular bands in the semigroup literature; most of the examples considered above are left regular bands. Under this condition the chain *is* diagonalizable. In later work, Brown showed that nearly everything goes through in the general case. There is a tantalizing extension to a walk on the chambers of a building. Here, while a product is well defined, it is not associative. This creates a mess but there are some positive results as well.

5. Some Open Problems

1. Almost none of the walks presented here have good lower bounds available. Examples include riffle shuffles with the deck cut exactly in two (see Section 3A) or any of Fulman's shuffles (Section 3C). It would be nice to have a lower bound in some generality for the general hyperplane walks of Section 4A. Usually, reasonable lower bounds are easier to prove than upper bounds. See [Diaconis, 1988] or [Saloff-Coste 1997] for the usual techniques. One idea for a systematic approach: Brown's method (Section 4C) finds a representation theoretic interpretation. With characters available, perhaps David Wilson's [2001] approach may be pushed through.

2. It should be the case that essentially all the walks discussed here show a sharp cutoff in their approach to stationarity; proving this requires sharp upper bounds as well as sharp lower bounds. The general upper bounds (e.g., Theorem 2 of Section 4A) are often slightly off in the few cases where sharp answers are known. For example, for ordinary riffle shuffles, the general approach shows $2\log_2 n + C$ shuffles suffice for randomness while Theorem 1 of Section 2A shows the right answer is $\frac{3}{2}\log_2 n + C$. The original paper of Bidigare-Hanlon-Rockmore gives a potentially sharper upper bound. It would be very instructive to compare the two variations. In preliminary work, Brown-Diaconis [1998] found them similar but Uyemura-Reyes [2002] found examples where the BHR bound is a genuine improvement. It may be that the bounds of BHR or Theorem 2 of Section 4A are sharp for some other metric; this happens for ordinary riffle shuffles with separation distance as discussed in Section 2C. At a more abstract level, it may be possible to prove

the existence of a sharp threshold without being able to locate it along the line of concentration inequalities see Ledoux [2000].

3. A very clear set of problems is to give any kind of upper bounds for Lalley's f-shuffles of Section 3C. Presumably, these all mix n cards in order $\log n$ steps but at present we don't know that order 2^n steps suffice.

4. For practical reasons it is natural to seek models of riffle shuffling cards that result in neater shuffles than the GSR shuffles. This arises in studying the way Las Vegas dealers shuffle; they drop cards in close to perfect alternation while the GSR method has packet sizes geometrically distributed. Here is a suggestion whose analysis is completely open: the *Markovian Model* is driven by a 2-state Markov chain with transition matrix.

$$\begin{pmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{pmatrix}$$

To shuffle a deck of n cards, run the chain starting in stationarity to produce $x_1, x_2, \ldots, x_n$ a binary sequence. If this sequence has k zeros and $n-k$ ones, cut off the top k cards as a left hand pile, the $n-k$ remaining as a right hand pile. Use the zeros and ones (from right to left say) to dictate if the next drop is from left or right.

For example, with $n = 10$ cards the sequence 0101100110 results in 5 cards being cut off and the final arrangement $1, 6, 2, 7, 8, 3, 4, 9, 10, 5$. This includes the GSR model by taking $p_{ij} = 1/2$ and perfect shuffles by taking $p_{01} = p_{10} = 1$. It is natural to begin with a symmetric cut, so $p_{01} = p_{10}$ and $p_{00} = p_{11}$. There is every hope that this model will produce neat and useful analyses. It must be the case that (for symmetric shuffles with $0 < p_{01} < 1$) there is a sharp threshold at $\theta \log n + C$ with $\theta = \theta(p)$. For practical purposes one could estimate $\theta(p)$ from computer experiments and also estimate p by watching dealers shuffle. This would allow one to derive reasonable ways of exploiting the structure if the dealers do not shuffle enough. The following seems clear: Since $p_{ij} = 1/2$ requires seven shuffles and this is the fastest method, most neat shufflers will require a good many more shuffles and there will be plenty of structure to take advantage of! Incidentally, the case of 'random perfect shuffles', each time choosing randomly to do a perfect in or out shuffle, has been analyzed by Uyemura-Reyes (2002). For perfect shuffles, not all permutations are possible so the walk is random on a subgroup. For decks of size 2^k, he shows order k^2 steps are necessary and suffice. The case of general decks remains open.

5. I want to record empirical work which yields a quite different natural model for riffle shuffling. In joint work with my student Arnab Chakraborty we studied commonly available machines for shuffling cards. These machines have the user cut the deck in two halves (placed into the left and right sides of the machine). Then a button is pushed which activates rubber wheels touching the bottoms of the two packets. These spin cards off the bottoms into a central region where they drop onto a collecting place. At the end of one shuffle the user retrieves the deck, cuts it into two and the process continues.

In our empirical work we found that an "opposite" GSR model seemed to fit the data. In the GSR model, if at some stage there are A cards in the left half and B cards in the right half, the chance of dropping the next card from the left half is $A/(A+B)$. In the mechanical shuffler, the chance seemed to be $B/(A+B)$; it was more likely for a card to be dropped from the smaller half. Of course, once all the cards in a half are used up, the remaining cards are dropped on top. This seems like a natural candidate for careful study.

One may interpolate between models with a one-parameter family of the following form. If at some stage there are A cards in the left half and B cards in the right half, the chance of dropping the next card from the left is $A^\theta B^{1-\theta}/(A^\theta B^{1-\theta} + B^\theta A^{1-\theta})$. Again, for practical purposes θ could be fit from data and a cutoff parameter could be estimated by simulation.

6. The GSR is the most uniform single riffle shuffle in the sense that, given the cut, it makes all shuffles equally likely. It is not clear (though it seems plausible) that it is also the probability on shuffles so that $Q * Q$ is closest to uniform (as well as Q^{*k} for all k). This may be a simple problem but it seems worth clarifying. A similar case that would shed some light: on $\mathbf{Z}/m\mathbf{Z}$ (the integers mod m), consider all probability measures with support in $[-a, a]$. Find the probability P in this set such that P^{*k} is closest to uniform (say in entropy or total variation distance). Is this uniform?

7. A beautiful set of conjectures has arisen from thesis work of J.C. Uyemura-Reyes. To describe them, consider first the random to top shuffle. This has eigenvalues $0, 1/n, 2/n, \ldots (n-2)/n$, 1 with multiplicity of j/n the number of permutations in S_n with j fixed points. This was proved by Phatarphod and independently by Wallach (1986) and follows from Theorem 1 of Section 4A. Next consider the multiplicative reversibilization of random to top. This is random to top followed by top to random: It may also be described as: remove a random card and insert it in a random position.

   Numerical work shows that the eigenvalues are all of the form quadratic function of $(j)/n^2$. For some cases this can be proved. For example, zero occurs with multiplicity the number of derrangements and the eigenvalues in representations near the trivial representation (or alternating representation) can be proved of this form. There must be a way to understand these! Work of Phil Hanlon and Patricia Hersh indicates that this question fits very neatly into algebra, along the lines of Section 3A.

   Using the available results one may conjecture where the cutoff occurs for mixing. For either random to top or top to random, $n \log n$ is the cutoff. It seems that random to random must be faster but perhaps not by more than a factor of two. We conjecture that $3/4n \log n$ is the cutoff here. This is what is required to kill the eigenvalues from the $n-1$ dimensional representation. Uyemura-Reyes [2002] proves a lower bound of form $\frac{1}{2} n \log n$ and an upper bound of $4n \log n$. At present writing we do not know that $n^2\times$ eigenvalue is an integer.

   Again, in preliminary work, it seems as if the eigenvalues of the multiplicative symmetrization of any hyperplane walk from a Coxeter group with symmetric face weights generated by a finite reflection group will be "nice" in the same

sense. To be specific, consider the permutation group S_n. Fix a composition $\mu = (\mu_1, \mu_2, \ldots, \mu_r)$ of n. A symmetric μ shuffle removes a uniformly chosen subset of μ_1 cards (keeping them in their same relative order, then, from the remaining $n - \mu_1$ cards, a random subset of size μ_2, and so on. With a final packet of size μ_r. These r packets are shuffled together by a GSR shuffle.

8. A very simple to state conjecture: After k GSR shuffles of n cards consider turning up cards from the top one at a time. What is the optimal guessing strategy to maximize the expected number of correct guesses? A conjectured optimal strategy due to McGrath is described in Bayer-Diaconis [1992]. There is related work in Ciucu (1998). Prove that McGrath's strategy is optimal. An easier version (still open) asks the same question following k top to random shuffles with k fixed and known.

9. Less of a conjecture than a suggestion; many of the semigroup walks in Brown [2000] seem worthwhile studying in depth. To take one example; Brown [2000, Section 6.3] introduced a fascinating family of walks on phylogenetic trees. Walks on such trees are currently an active area of study. See Diaconis-Holmes [2002] for pointers to work by Aldous and to the currently very active work in biology. Brown's walks are driven by weights w_{ij}. A first natural problem is to study the stationary distribution of Brown's walk as a natural family of non-uniform distributions on trees. They carry over to trees the Luce model which has been very actively studied for permutations. One might even contemplate estimating Brown's parameters w_{ij} from data. It is also natural to carry out some careful analyses of rates of convergence for natural families of weights: randomly chosen i.i.d. uniform weights, Zipf type weights, or w_{ij} the distance from i to j in some natural geometric structure.

10. A most annoying problem: find some use for the eigenvalues of the many walks in the section above. For reversible Markov chains there are good bounds on the rate of convergence based on eigenvalues. Are there *any* explicit bounds on, e.g., L^2 distances for non-symmetric chains? Going further, are there bounds on the multiplicative symmetrization of a chain based on knowledge of the eigenvalues of the original chain? This would allow the wealth of eigenvalue information reported above to be used for comparison purposes as explained in Saloff-Coste [1997].

REFERENCES

Aguiar, M. and Sottile, F. (2002), Structure of the Malvenuto-Reutenauer Hopf algebra of permutations. Technical Report, Dept. of Mathematics, University of Massachusetts, Amherst.

Aldous, D. (1983), Random walks on finite groups and rapidly mixing Markov chains, in *Springer Lecture Notes in Mathematics* **986**.

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

Aldous, D. and Fill, J. (2002, *Reversible Markov Chains*, Available at http://www.stat.berkeley.edu/users/aldous.

Bayer, D. and Diaconis, P. (1992), Trailing the Dovetail shuffle to its lair, *Ann. Appl. Probab* **2**, 294-313.

Berger, P. (1973), On the distribution of hand patterns in bridge: Man-dealt versus computer dealt, *Canadian Jour. Statist.* **1**, 261-266.

Bergeron, F., Bergeron, W. and Garsia, H. (1989), *The Free Lie Algebra and q-Enumeration.* In D. Stanton (ed.), *Invariant Theory and Tableaux*, Springer, New York, 166-190.

Bergeron, W. (1991), A hyperoctahedral analogue of the free Lie algebra, *Jour. Combin. Th. A.* **55**, 80-92.

Bergeron, F. and Bergeron, N. (1992), Orthogonal idempotents in the descent algebra of B_n and applications, *Jour. Pure Appl. Alg.* **79**, 109-129.

Bergeron, F., Bergeron, N., Howlett, R. and Taylor, D. (1992), A decomposition of the descent algebra of a finite Coxeter group, *Jour. Alg. Combinatorics* **1**, 23-44.

Bergeron, N. (1994), *On Hochschild homology and Vassikiev invariants*, DIMACS Conference, 1-10.

Bergeron, N. (1995a), Hyperoctahedral operations on Hochschild homology, *Adv. Math.* **110**, 255-276.

Bergeron, N. (1995b), Décomposition hyperoctahédrale de l'homologie de Hochschild, *Discrete Math.* **139**, 33-48.

Bergeron, W. and Wolfgang, H. (1995), The decomposition of Hochschild cohomology and Gerstenhaber operations, *Jour. Pure Appl. Alg.* **104**, 243-265.

Bidigare, P. (1997), Hyperplane arrangements face algebras and their associated Markov chains. Ph.D. thesis, University of MIchigan, Department of Mathematics.

Bidigare, P., Hanlon, P. and Rockmore, D. (1999), A combinatorial description of the spectrum for the Tsetlin library and its generalization to hyperplane arrangements, *Duke Math. Journal* **99**, 135-174.

Billera, L., Brown, K. and Diaconis, P. (1999), Random walks and plane arrangements in three dimensions, *Amer. Math. Monthly* **106**, 502-524.

Billera, L. Hsiao, S. and Van Willigenburg (2001), Peak quasi-symmetric functions and eulerian enumeration. Preprint, Dept. of Mathematics, Cornell University.

Billera, L. and Sarangarajan (1996), The combinatorics of permutation polytopes. In L. Billera et al. (eds.). *Formal Power Series and Algebraic Combinatorics*, Amer. Math. Soc. Providence.

Borel, E. and Chéron, A. (1940), Théorie mathématique du bridge a la portée de tous. Gauthier-Villars, Paris. English translation by Alec Traub, (N.D.), Monna Lisa Publishing, Taiwan.

Buhler, J., Eisenbud, D. Graham, R. and Wright, C. (1994), Juggling drops and descents, *Amer. Math. Monthly* **101**, 507-519.

Brown, K. (2000), Semigroups, rings, and Markov chains, *Jour. Theoretical Probability* **13**, 871-930.

Brown, K. (2001), Notes on bands. Preprint, Department of Mathematics, Cornell University.

Brown, K. and Diaconis, P. (1998), Random walks and hyperplane arrangements, *Ann. Probab.* **26**, 1813-1854.

Cartan, H. (1967), *Formes Différantielles*, Hermann, Paris.

Carter, R. (2002), Semisimple conjugacy classes and classes in the Weyl Group. Preprint Mathematics Institute, University of Warwick.

Ciucu, M. (1998), No-feedback card guessing for dovetail shuffles, *Ann. Appl. Probab.* **8**, 1251-1269.

Diaconis, P. (1988), *Group Representations in Probability and Statistics*, IMS, Hayward, CA.

Diaconis, P. (1996), The cutoff phenomenon in finite Markov chains, *Proc. Nat. Acad. Sci. USA* **93**, 1659-1664.

Diaconis, P. (1998), From shuffling cards to walking around the building: An introduction to modern Markov chain theory. In *Proc. Int. Congress Math.* **1**, 187-204.

Diaconis, P., Graham, R. and Kantor, W. (1983), The mathematics of perfect shuffles, *Adv. Appl. Math.* **4**, 175-193.

Diaconis, P., Fill, J. and Pitman, J. (1992), Analysis of top to random shuffles, *Combinatorics Probability and Computing* **1**, 135-155.

Diaconis, P., McGrath, M. and Pitman, J. (1995), Riffle shuffles, cycles and descents, *Combinatorica* **15**, 11-20.

Diaconis, P. and Holmes, S. (2000), Analysis of a card mixing scheme. Unpublished report.

Diaconis, P. and Holmes, S. (2001), Random walk on trees and matchings, *Electronic Jour. Probab.* **7**.

Dobrow, R. and Fill, J. (1995), *The Move to the Front Rule for Self-organizing Lists with Markov Dependence*. In D. Aldous et al. Ed. *Discrete Probability and Algorithms*, Springer, New York, pg. 51-80.

Doob, J. (1954), *Stochastic Processes*, Wiley, N.Y.

Dupont, J. (2001), *Scissors Congruences, Group Homology and Characteristic Classes*, World Scientific, Hong Kong.

Fill, J. (1996), An Exact formula for the move-to-front rule for self-organizing lisis, *Jour. Theoret. Prob.* **9**, 113-160.

Fulman, J. (1998), The combinatorics of biased riffle shuffles, *Combinatorics* **18**, 173-174.

Fulman, J. (1999), Counting semisimple orbits of finite Lie algebras by genus, *Jour. Algebra* **217**, 170-179.

Fulman, J. (2000A), Semisimple orbits of Lie algebras and card-shuffling measures on Coxeter groups, *Jour. Algebra* **224**, 151-165.

Fulman, J. (2000B), Affine shuffles, shuffles with cuts, the Whitehouse module, and patience sorting, *Jour. Algebra* **231**, 614-639.

Fulman, J. (2001A), Descent algebras, hyperplane arrangements, and shuffling cards, *Proc. Amer. Math. Soc.* **129**, 965-973.

Fulman, J. (2001B), Applications of the Brauer complex: Card shuffling, permutation statistics, and dynamical systems, *Jour. Algebra* **243**, 96-122.

Fulman, J. (2002), Applications of symmetric functions to cycle and increasing subsequence structure after shuffles. Preprint, Department of Mathematics, University of Pittsburgh. To appear, *Jour. Alg. Combinatorics*.

Fulton, W. (1997), *Young Tableaux*, Cambridge Press, Cambridge.

Garsia, A. and Remmel, J. (1985), Shuffles of permutations and the Kronecker product., *Graphs Combinatorics* **1**, 217-263.

Garsia, A. and Reutenauer, C. (1989), A decomposition of Solomon's descent algebra, *Adv. in Math.* **77**, 189-262.

Garsia, A. (1990), Combinatorics of the free Lie algebra and the symmetric group. In *Analysis ETC*, Jurgen Moser Festschrift. Academic Press, N.Y., pg. 309-82.

Garsia, A. (2002), *On the powers of top to random shuffling.* Typed notes, UCSD.

Garsia, A. and Wallach, N. (2002), Quasi-symmetric functions modulo symmetric functions are Cohen-Macually. Preprint, Dept. of Mathematics, UCSD.

Gerstenhaber, M. and Schack, S. (1987), A hodge-type decomposition for commutative algebra cohomology, *Jour. Pure Apl. Alg.* **48**, 229-247.

Gilbert, E. (1955), Theory of Shuffling. Technical Report, Bell Laboratories.

Gessel, I. and Reutenauer, C. (1993), Counting permutations with given cycle structure and descent set, *Jour. Combin. Theory. Ser. A* **64**, 189-215.

Green, J. and Rees, D. (1952), On semigroups in which $x^r = x$, *Proc. Camb. Phil. Soc.* **48**, 35-40.

Greenbaum, A. (2002), Card shuffling and the polynomial numerical hull of degree k. To appear, *SIAM J. Sci. Comput.*

Hanlon, P. (1990), The action of S_n on the components of the Hodge decomposition of Hochschild homology, *Mich. Math. Jour.* **37**, 105-124.

Hanlon, P. (1992), Order and disorder in algebraic combinatorics, *Math. Intell* **14**, 20-25.

Hanlon, P. and Hersh, P. (2002), A hodge decomposition for the complex of injective words. Technical Report, Dept. of Mathematics, University of Michigan.

Hatcher, A. (2002), *Algebraic Topology*, Cambridge University Press, Cambridge.

Jonsson, G. and Trefehten, L. (1998), A numerical analyst looks at the cutoff phenomenon in card shuffling and other Markov chains, in *Numerical Analysis* **1997**, (D. Griffiths et al. (eds.), Addison Wesley.

Keller, J. (1995), How many shuffles to mix a deck? *SIAM Rev* **37**, 88-89.

Lalley, S. (1996), Cycle structure of riffle shuffles, *Ann. Probab.* **24**, 49-73.

Lalley, S. (1999), k-Riffle shuffles and their associated dynamical systems, *Jour. Theoret. Probab.* **12**, 903-932.

Lalley, S. (2000), On the rate of mixing for p-shuffles, *Ann. Appl. Prob.* **10**, 1302-1321.

Ledoux, M. (2000), *Concentration Measures*, American Math. Soc. Providence.

Loday, J. (1988), Partition Eulerienne et operations en homologie cyclique, *C.R. Acad. Sci. Paris Ser 1 Math.* **307**, 283-286.

Lovasz, L. and Winkler, P. (1995), Mixing of random walks and other diffusions on a graph. In *Surveys in Combinatorics* (P. Rowlinson ed.) London Math Soc. Lecture Notes, V. 218, pg. 119-154, Cambridge University Press, Cambridge.

Mann, B. (1995), How many times should you shuffle a deck of cards? In *Topics in Contemporary Probability and its Applications* (J.L. Snell, ed.) CRC Press, Boca Raton.

Macdonald, I. (1985), *Symmetric Functions and Hall Polynomials*, Clarendon Press, Oxford.

MacLane, S. (1950), Cohomology of Abelian groups, *International Congress of Mathematicians* **2**, 8-14.

MacLane, S. (1963), *Homology*, Springer, Heidelberg.

MacLane, S. (1970), The Milgram bar construction as a tensor product of functions, Springer Lecture Notes in Math, No. 168, pg. 135-152.

Mahajan, S. (2002), Shuffles on Coxeter groups. Preprint, Dept. of Math., Cornell University.

Markov, A. (1906), Extension of the law of large numbers to dependent events (Russian), *Bull. Soc. Math. Kazan* **2**, 155-156.

Milgram, J. (1967), The bar construction and abelian H-spaces, *Ill. Jour. Math.* **11**, 242-250.

Morris, S.B. (1998), *Magic Tricks, Card Shuffling and Dynamic Computer Memories*, M.A.A. Washington, D.C.

Pemantle, R. (1989), An analysis of the overhand shuffle, *Jour. Theoret. Probab.* **2**, 37-50.

Phatarfod, R. (1991), On the matrix occuring in a linear search problem, *Jour. Appl. Probab.* **28**, 336-346.

Poincare, H. (1912), Calcul des probabilités, 2nd ed. Gauthier Villars, Paris.

Reeds, J. (1981), Theory of shuffling, unpublished manuscript.

Reutenauer, C. (1993), *Free Lie Algebras*, Clarendon Press, Oxford.

Saloff-Coste, L. (1997), Lectures on finite Markov chains, *Springer Lecture Notes in Math* **1665**, 301-408.

Saloff-Coste, L. (2001), Probability on groups: Random walks and invariant diffusions, *Notices Amer. Math. Soc.* **48**, 968-977.

Saloff-Coste, L. (2002), Random walks on finite groups in H. Kesten (ed.), *Springer Encyclopedia of Mathematical Sciences, Discrete Probability Volume.* To appear.

Shahshahani, M. (2002), *Lecture Notes on Geometry*,

Shnider, S. and Sternberg, S. (1993), *Quantum Groups*, International Press, Cambridge, MA.

Solomon, S. (1976), A Mackey formula in the group ring of a Coxeter group, *Journal of Algebra* **41**, 255-268.

Stadler, J. (1997), Schur functions, juggling, and statistics on shuffled permutations, Ph.D. Dissertation, Dept. of Mathematics, Ohio State.

Stanley, R. (1972), *Ordered Structures and Partitions*, Memoirs Amer. Math Soc. **119**, Amer. Math Soc. Providence.

Stanley, R. (1997), *Enumerative Combinatorics*, Vol. I, 2nd ed., Cambridge University Press, Cambridge.

Stanley, R. (1999), *Enumerative Combinatorics*, Vol. II, Cambridge University Press, Cambridge.

Stanley, R. (2002), Generalized riffle shuffles and quasisymmetric functions. *Annals of Combinatorics* **5**, 479-491.

Stark, D., Ganesh, D. and O'Connell, N. (2002), Information loss in riffle-shuffling, *Combin. Probab. Comput.* **11**, 79-95.

Su, F. (1995), Methods for quantifying rates of convergence for random walks on groups. Ph.D. Thesis, Harvard University.

Steenrod, N. (1967), Milgrams classifying space of a topological group, *Topology* **7**, 319-368.

Thorpe, E., (1972), Non-random shuffling with applications to the game of faro, *Jour. Amer. Statist. Assoc.* **68**, 842-847.

Trefethen, L. and Trefethen, L. (2000), How many shuffles to randomize a deck of cards? *Proc. Roy. Soc. London A* **456**, 2561-2568.

Uyemura-Reyes, J.C. (2002), Random walk, semi-direct products, and card shuffling, Ph.D. Thesis, Dept. of Mathematics, Stanford University.

Wallach, N. (1988), Lie algebra cohomology and holomorphic continuation of generalized Jacquet integrals, *Adv. Studies Pure Math.* **14**. Representations of Lie groups, Hiroshima, 123-151.

Wilson, D. (2001), Mixing times of lozenge tiling and card shuffling. Preprint Microsoft Research, Seattle. To appear, *Ann. Appl. Probab.*