CHAPTER 1

Shuffling cards: An introduction

This chapter gives a sampling of the kind of problem we will analyze in this book and a brief description of what is covered in each chapter. It concludes with references to needed mathematical background.

1.1. Riffle shuffling and total variation

While there are all kinds of shuffles discussed in this book, the most common one we will refer to is the usual method of riffle shuffling: take a deck of $n$ cards, cut it “about in half”, and “riffle the two halves together”. This is repeated several times. One natural question: “How many times should the cards be shuffled so that they are thoroughly mixed?”

We begin by defining our terms. Throughout, picture a deck of $n$ cards, labeled $1, 2, \ldots, n$, starting in order with card 1 on top.

**Cut:** Cut the deck into two piles, cutting off $c$ cards with probability \( \binom{n}{c}/2^n \). This is the discrete version of the “bell-shaped curve”. It is a mathematical model for cutting the deck “about in half”.

**Riffle:** The two halves are now riffled together according to the “Gilbert-Shannon-Reeds model” (GSR model for short). At any stage there are two piles, say a left-hand pile containing $A$ cards and a right-hand pile containing $B$ cards. Drop the next card from the bottom of the left pile with probability $A/(A+B)$, and from the bottom of the right pile with probability $B/(A+B)$. Continue sequentially until all the cards have been dropped. The two interlaced piles are now pushed together:

\[
\begin{array}{cccccccc}
1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\
\text{CUT} & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\
1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\
\end{array}
\]

For example, suppose that $n = 52$ and that initially 26 cards are cut off. The first card is dropped from the left with probability $26/52 = 1/2$. Suppose that left is dropped first. The left pile now has 25 cards. The chance that the next is dropped from the left is thus $25/51$, slightly less than $1/2$. This is continued until all the cards are dropped.

The GSR model is physically intuitive (if cards are dropped by “springyness”, they are more likely to be dropped from the bigger pile). Experiments with real people shuffling real cards (discussed in Section 4D of [142]) show that it is a good
model for the way normal people shuffle. It is not a good model for the way that
Vegas dealers (or one of your authors) shuffle. They are able to shuffle close to
perfectly (left, right, left, right, ...), and we remark that if the cards are perfectly
interlaced, then eight perfect shuffles bring the deck back to its original order (see
Section 7.1).

Shuffling twice: Let \( \pi \) be a permutation of \( 1, 2, \ldots, n \). We think of \( \pi \) as an
arrangement of a deck of cards with \( \pi(1) \) the label of the top card, \ldots, and \( \pi(n) \) the
label of the bottom card. Let the chance of being at \( \pi \) following one GSR shuffle
be \( Q(\pi) \). The chance of \( \pi \) following two shuffles is denoted \( Q^*2(\pi) \). By definition,
this is
\[
Q^*2(\pi) = \sum_\eta Q(\eta)Q(\pi \eta^{-1}).
\]
Thus, to wind up at \( \pi \) after two shuffles, we have to choose some arrangement \( \eta \)
for the first shuffle and then choose the permutation that goes from \( \eta \) to \( \pi \), namely
\( \pi \eta^{-1} \), for the second shuffle. In our model the two choices are independent. More
complicated choices are possible: see the discussion in Section 8.6.

\( k \)-shuffles: The chance of achieving a given permutation \( \pi \) after \( k \) shuffles is
given by the formula
\[
Q^*k(\pi) = \sum_\eta Q(\eta)Q^*(k-1)(\pi \eta^{-1}).
\]
For completeness, \( Q^0(\pi) \) is 1 if \( \pi \) is the original arrangement \( 1, 2, \ldots, n \) and 0 for
other \( \pi \). And \( Q^*1(\pi) = Q(\pi) \).

Quantifying mixing. The uniform distribution assigns equal probability \( U(\pi) = 1/n! \) of achieving any given permutation \( \pi \). The aim is to make all possible ar-
rangements equally likely. Alas, no finite number of riffle shuffles leads to perfect
uniformity. There is always a bias. To quantify this we need a way to measure
what it means to be “close to uniform”.

There are many possible definitions of “distance to uniform”. A basic one is
total variation:
\[
\|Q^k - U\|_{TV} = \max_A |Q^k(A) - U(A)|.
\]
Here \( A \) ranges over all subsets of arrangements. For example, \( A \) might be the set
of arrangements with the ace of spades in the top half of the deck. The notation is
\[
Q^k(A) = \sum_{\pi \in A} Q^k(\pi).
\]
That is, \( Q^k(A) \) is the chance that the deck is in one of the arrangements in \( A \) after
\( k \) shuffles. Thus if \( \|Q^k - U\|_{TV} \) is small (less than 1/1,000, say), then \( Q^k(A) \) and
\( U(A) \) are close, no matter what set \( A \) is being considered. Many other distances
are used (and discussed below; see Section 3.7). Total variation is a gold standard.
1.2. The problem, its motivation, and some theorems

With the above preparatory definitions, we are now ready to transform the question of, “How far is a given shuffling technique from truly randomizing a deck of cards?” into a well-posed mathematical problem: given $\epsilon > 0$, how large should $k$ be so that

$$\|Q^k - U\|_{TV} < \epsilon$$

Precise answers to this question are described in this book (when $n = 52$, as we explain just after Theorem 1.2.1 about 7 shuffles suffice).

Before considering this problem, let us discuss some reasons both practical and mathematical for studying how effective shuffling is at randomizing cards.

(1) How people shuffle does matter. In friendly family games, poker tournaments, casinos, and bridge clubs, cards are shuffled and the game depends on it. To see that shuffling matters, consider a natural experiment performed for us when the American Bridge League switched from hand shuffling to computer shuffling. There are records of 1,000 tournament deals before and after the switch. Recall that in bridge fifty-two cards are dealt to four hands of thirteen each. The suit distribution (number of clubs, diamonds, hearts, and spades) in each hand is important. The data below gives the suit distribution for one of the hands (South). Each deal results in four numbers, e.g., $4, 3, 3, 3$ — the number of clubs, diamonds, hearts, and spades in the deal. If these are “flat” (close to even, as $4, 3, 3, 3$), the hand is unexciting. If they are skewed, as $0, 11, 2, 0$, the hand will be interesting. Bridge players had noticed that after the computer started determining the hands, many more skewed hands appeared. They complained that the computer was faulty. In fact as we are about to see, the opposite is true, and mathematics helped to change the way professionals shuffle cards.

Table 1.1 gives data taken from an article of Berger [53]. See also [84].

<table>
<thead>
<tr>
<th>Distribution of the 4 suits</th>
<th>Column 1</th>
<th>Column 2</th>
<th>Column 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>$4, 4, 3, 2$</td>
<td>216</td>
<td>198</td>
<td>241</td>
</tr>
<tr>
<td>$5, 3, 3, 2$</td>
<td>155</td>
<td>160</td>
<td>172</td>
</tr>
<tr>
<td>$5, 4, 3, 1$</td>
<td>129</td>
<td>116</td>
<td>124</td>
</tr>
<tr>
<td>$5, 4, 2, 2$</td>
<td>106</td>
<td>92</td>
<td>105</td>
</tr>
<tr>
<td>$4, 3, 3, 3$</td>
<td>105</td>
<td>103</td>
<td>129</td>
</tr>
<tr>
<td>$6, 3, 2, 2$</td>
<td>56</td>
<td>64</td>
<td>46</td>
</tr>
<tr>
<td>$6, 4, 2, 1$</td>
<td>47</td>
<td>53</td>
<td>36</td>
</tr>
<tr>
<td>$6, 3, 3, 1$</td>
<td>34</td>
<td>40</td>
<td>41</td>
</tr>
<tr>
<td>$5, 5, 2, 1$</td>
<td>32</td>
<td>40</td>
<td>19</td>
</tr>
<tr>
<td>$4, 4, 4, 1$</td>
<td>30</td>
<td>35</td>
<td>25</td>
</tr>
<tr>
<td>7, 3, 2, 1 and others</td>
<td>90</td>
<td>99</td>
<td>62</td>
</tr>
</tbody>
</table>

1,000 1,000 1,000

Table 1.1. Column 1: expected frequencies, Column 2: actual frequencies of computer dealt hands, Column 3: actual frequencies of man-dealt hands.
A standard chi-square test comparing Column 2 (observed) with Column 1 (expected) is 10.31 on ten degrees of freedom. We expect a chi-square variate on ten degrees of freedom to be about 10, so there is no significant difference between the computer-dealt hands and chance. On the other hand, the chi-square statistic comparing Column 3 with Column 1 is 31.1, a wildly significant difference between hand-shuffled deals and what is expected by chance.

Faulty shuffling, in-the-hand shuffled era, resulted in too many flat hands. To understand a natural mechanism for this, recall that in bridge, if the lead player plays clubs, all the players must play clubs (if possible), and whoever plays the highest card gathers the four cards into a packet. At the end of a deal, the thirteen packets are gathered together into one. Thus the cards tend to be in groups of four of the same suit. Now picture a few lazy shuffles. Cards which started together tend to stay together. When cards are dealt for the next hand \(1, 2, 3, 4, \ldots\), the clumps are evenly distributed and the resulting distribution is flat. The data is unmistakeable. The computer-generated deals pass all standard tests for randomness.

(2) Shuffling is a basic mathematical operation. Represent a pile of cards as a sequence of letters. For example \(AB\) is a pile of two cards with \(A\) on top and \(B\) on the bottom. We shall use the symbol \(AB\omega XYZ\) to stand for the sum of the possible interleavings of \(AB\) with \(XYZ\), which keep \(AB\) and \(XYZ\) in the same relative order. For example one such interleaving is \(AXBYZ\), and taking the sum of all possible interleavings gives

\[
ABXYZ + AXYBZ + AYBZ + YABZ + YAZB + XYZA.
\]

The patterns of shuffling can be expressed in the language of multilinear functions and differential forms. By making these analogies mathematically precise, it is possible to use our hands-on understanding of shuffling to give simple proofs of classical facts in mathematics and even come up with new results. Conversely, the power of mathematics developed over the millennia can be harnessed to give useful new insights and results about shuffling.

We illustrate these connections throughout the book. See, for example, Chapter 10 for topology and Chapters 6 and 12 for combinatorics.

(3) Shuffling illuminates “Markov chains”. Almost every area of our world can be studied by simulation. A widely used method of simulation is called Markov chain Monte Carlo. One runs an iterative process, akin to repeated shuffling, to wind up with a desired distribution (the uniform distribution for shuffling). The study of how long such algorithms need to be run to do their job has shuffling as a special case. By luck and hard work, precise answers are available for shuffling. These allow proofs of phenomena that hold much more generally, such as rapid mixing and the cutoff phenomenon.\[320\]
Some theorems about card shuffling: What does a theorem in card shuffling look like? Here are two examples proved by Bayer and Diaconis [46]. The first result is for $n = 52$ with exact values. The second result is for general $n$.

**Theorem 1.2.1.** Let $Q^k(\pi)$ be the chance that a 52-card deck is in order $\pi$ after $k$ riffle shuffles. Table 1.2 gives the distance to uniformity.

**Table 1.2. Total variation distances after $k$ riffle shuffles.**

<table>
<thead>
<tr>
<th>$k$</th>
<th>$1$</th>
<th>$2$</th>
<th>$3$</th>
<th>$4$</th>
<th>$5$</th>
<th>$6$</th>
<th>$7$</th>
<th>$8$</th>
<th>$9$</th>
<th>$10$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$|P^k - U|_{TV}$</td>
<td>$1.000$</td>
<td>$1.000$</td>
<td>$1.000$</td>
<td>$1.000$</td>
<td>$.924$</td>
<td>$.614$</td>
<td>$.334$</td>
<td>$.167$</td>
<td>$.085$</td>
<td>$.043$</td>
</tr>
</tbody>
</table>

The distance $\|Q^k - U\|_{TV} = \max_A |Q^k(A) - U(A)|$ is the difference between two numbers between zero and one, so it is between zero and one. It is one if $Q^k$ and $U$ are completely different (supported on disjoint sets). It is zero if $Q^k$ and $U$ are exactly the same.

Theorem [1.2.1] is sometimes called the “seven shuffles theorem”. This makes referring to it easy, but what do we mean by this name? From the numbers, the distance to uniformity stays at its maximum until five or so shuffles when it begins to descend to zero. There is a sharp cutoff from 5 to 9 shuffles, when it tends to zero exponentially fast. The number “7” is in the middle of this cutoff, after which the distance falls by roughly a factor of two after each shuffle. The theory shows this continues forever. Of course, practical application is motivated by circumstances. If you are playing “Go Fish” with your kids or grandparents, shuffle 5 to 6 times. If national security depends on it, shuffle 11 times.

The next theorem applies to deck sizes of arbitrary length.

**Theorem 1.2.2.** If $n$ cards are riffle shuffled $k$ times with $k = \frac{3}{2} \log_2(n) + c$, then for large $n$,

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

with

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

To help parse this, consider $k$ shuffles of $n$ cards. These determine $c = c(n,k)$ from $k = \frac{3}{2} \log_2(n) + c$. This is the same $c$ appearing on the right side of the equation. The behavior of $\Phi(x)$, the normal distribution, is very well understood. This shows that the total variation distance tends exponentially fast to 0 as $c$ tends to infinity (the powers of 2 in Table 1.2). It also shows that the total variation distance tends to 1 doubly exponentially fast if $c$ tends to negative infinity.

The theorem gives the asymptotic shape of the transition from order to fully mixed. It shows that there is a sharp cutoff at $\frac{3}{2} \log_2(n)$. The numbers in Table 1.2 show that these asymptotics are useful for deck sizes of interest.

**1.3. Outline of the book**

We will prove theorems such as Theorems 1.2.1 and 1.2.2. Along the way we link the mathematics of shuffling to other parts of mathematics and to the real world of gambling and games.
Chapter 2 illustrates various types of shuffles studied in this book. It also gives a history of shuffling.

Chapter 3 sets the stage, defining a variety of distances (how far from random are we?), proving Theorem 1.2.1 along the way. This includes separation distance, total variation distance, and relative entropy. It tells what we know about the effects of cuts. Finally, some ways of taking advantage of badly shuffled cards are developed.

Chapter 4 studies the behavior of various features of the deck after shuffling. After all, we may not care about the complete arrangement. In blackjack the suits don’t matter and 10, J, Q, K all count the same. Fewer shuffles may suffice. One highlight is a (provable) “rule of thumb”: for most features, \( \log_2(n) \) shuffles suffice. For some features there are nice formulas that are connected to other parts of mathematics such as algebra and Lie theory. We also describe the effect of different methods of dealing, which can significantly enhance randomness.

Chapter 5 treats shuffling as a Markov chain, here a Markov chain on the space of all \( n! \) permutations. The eigenvalues and eigenvectors of this Markov chain can be expressed in “closed form”. This is closely related to a seemingly far afield realm of mathematics: the eigenvalues link to “Hodge decompositions of Hochschild homology” and the eigenvectors link to Hopf algebras. An appendix to this chapter gives a list of applications of eigenvalues and eigenvectors to Markov chains.

Chapter 6 goes in a completely different direction, relating shuffling to the behavior of “carries” in ordinary arithmetic. Consider adding two large integers, say

\[
\begin{array}{c|c|c|c|c|c}
& 1 & 11011 & 10110 \\
& 49567 & 75326 & & & \\
& 86614 & 59275 & & & \\
\hline
1 & 36182 & 34601 & & & \\
\end{array}
\]

The 1’s along the top are “carries” (here 8 carries and 3 “noncarries”). For large random numbers, it is natural to ask: “How many carries are there and how are they distributed?” Surprisingly, this is intimately connected with riffle shuffling! Successive carries form a Markov chain. When adding \( n \) large integers base \( b \), the carries Markov chain is exactly the same as the Markov chain of the number of descents when \( n \) cards are repeatedly “\( b \)-shuffled”. This means that the wonderful work on carries (work of John Holte \([269, 270]\)) illuminates shuffling and our shuffling work translates to answer natural questions about basic arithmetic.

Chapter 7 goes back to shuffling. It studies different models of shuffling such as the Thorp shuffle and a variety of variations using iterated maps. We also discuss the mathematics of perfect shuffles. The chapter concludes by considering shuffling of big decks.

Chapter 8 considers the move-to-front shuffle and some variations. It reviews the concept of coupling and gives a complete spectral analysis of the move-to-front shuffle. It relates the move-to-front shuffle with a statistical model of Plackett and Luce and gives a connection with “Stein’s method”.

Chapter 9 studies a vast generalization of shuffling via the geometry of hyperplane arrangements. Here a hyperplane arrangement is just a collection of (affine) hyperplanes in Euclidean space. Figure 1.1 gives an example of an arrangement of lines in the Euclidean plane.
1.3. OUTLINE OF THE BOOK

An arrangement divides space into chambers (the full-dimensional parts not on any of the hyperplanes) and faces (points on some of the hyperplanes and on one side of others). There is a natural projection operator which assigns to each chamber $C$ and each face $F$ the unique chamber $C'$ adjacent to $F$ and closest to $C$ (in the sense of crossing the fewest number of hyperplanes). See the example in Figure 1.1. With this set-up, a Markov chain (random walk on the chambers) can be defined by repeatedly choosing faces from a fixed probability distribution and moving from the current chamber $C$ to the chamber $C'$ adjacent to the chosen face and closest to $C$. Remarkably, this scheme includes riffle shuffling as a special case (the braid arrangement). It also includes a huge collection of other natural Markov chains such as the “move-to-front” chain of Chapter 8. Surprisingly, a fairly complete, unified analysis of all of these chains is available.

Chapter 10 considers connections between shuffling and algebraic topology. Standard theorems involving the topology (more precisely, homology and cohomology) of the product of two spaces in terms of the topology of the individual spaces and their intersection involve shuffling.

Chapter 11 studies Type B shuffles, also known as hyperoctahedral shuffles. These arose in our consulting work in the casino world while we looked at the behavior of card shuffling machines. The mathematics allowed us to give a thorough analysis of a real-world problem. The real-world problem suggested fresh mathematical questions (and so it goes).

Shuffling is a basic part of enumerative combinatorics. This subject is unified via symmetric function theory. There is a neat class of shuffles related to one of the most basic ingredients: Schur functions. Further, some important pillars of symmetric function theory (descent algebras, $P$-partitions, and quasisymmetric...
functions) are essentially the same as the mathematics of riffle shuffling. These topics are treated in Chapter 12.

Chapter 13 studies overhand shuffling. It gives a further introduction to coupling. It also gives two connections to higher mathematics: hyperplane arrangements on the one hand and ergodic theory, dynamical systems, and Riemann surfaces on the other.

Chapter 14 studies a common method of shuffling cards: “smoosh” shuffling. Here the cards are slid around “for a while” on the table before being gathered up and dealt. The mathematical techniques needed are quite different (stochastic calculus makes an appearance) but coupling is still useful in analyzing how long “for a while” needs to be to mix the cards.

Chapter 15 explains the standard method of shuffling cards on a computer. This is abstracted to the subgroup algorithm which is a practical method of generating random elements of any finite group. We then go “from algorithm to theorem”, showing how the factorization of the uniform distribution yields new “pure math” theorems.

Chapter 16 describes some applications of shuffling. This includes applications to magic tricks and shuffle tracking. It explains the first good card trick based on shuffling (Charles Jordan’s “Mindreading by Mail”) as well as the Gilbreath principle. There is a fascinating connection of the Gilbreath principle to the “Mandelbrot set” of dynamical systems theory (see Chapter 5 of the book [158]). Chapter 16 also explains “neocheating” techniques that take advantage of poor shuffling. We give some speculative connections with cars merging in traffic and statistics of permutations.

Chapter 17 describes the connection between multiple zeta values and shuffling. There are surprising equalities between the values of Riemann’s zeta function and “higher zeta values”. These are of interest in number theory, quantum field theory, and general mathematical curiosity. Conjecturally, all the relations can be explained in the language of shuffling.

1.4. Background for reading this book

Our analysis of shuffling has links to several parts of mathematics: to probability (in particular Markov chains), to combinatorics (enumerative, algebraic, and analytic), and to group theory. We have tried to make things self-contained (meaning that a math graduate student should find most of it accessible). Magicians might try reading Chapters 1, 2, 13, and 16. Additional pointers to background needed for special topics appear along the way. Here are some pointers to useful expository accounts of the relevant mathematics.

**Probability:** The introductory accounts by Ross [415], Grinstead and Snell [251], and particularly Feller (Volume 1) [194] cover the basics; Grinstead and Snell even cover riffle shuffling and we learned the mix of probability and analysis used throughout from Feller. A useful transition to fancier probability is in Lange [315]. A superb first graduate text is Billingsley [63]. A truly wonderful book on “rates of convergence for mixing” is by David Levin and Yuval Peres [320].

**Combinatorics:** For an introduction, try Bóna [71]. The more advanced treatment of enumerative and algebraic combinatorics in Stanley [438], [439] and
in Aigner [8] is inspiring. For analytic combinatorics, see Flajolet and Sedgewick [199], Melczer [346], or Pemantle and Wilson [382].

**Group theory:** Much of this book just uses the most basic properties of permutations as covered in any undergraduate group theory text [222]. For finer results (character theory of the symmetric group, tableaux combinatorics, etc.) see Sagan [416] or Macdonald [332].
CHAPTER 2

Practice and history of shuffling cards

While the introductory chapter focused on riffle shuffles, there are numerous other shuffles of practical interest. To help the reader picture what’s going on, some of these are illustrated in Section 2.1 with more careful definitions later in the book. Following this, in Section 2.2 we give some indication of the (early) history of shuffling cards.

2.1. Illustrations of some shuffles

This section illustrates various shuffle methods and is divided into eight sub-sections.

2.1.1. Riffle shuffling. The most common method of shuffling cards in Western societies is the riffle shuffle. This is performed in several ways.

In the hands: This is shown in Figure 2.1: the deck is held in the hands, about half is dropped into the other hand, the two packets are riffled together and then pushed together. This is repeated several times.

With a flourish: Some people spring the cards together. See Figure 2.2.

On the table: In casinos and card clubs where serious gambling occurs, people worry that the value of the bottom card may be exposed during the in-the-hands shuffle and the entire procedure is carried out with the cards on the table. See Figure 2.3.
Butt shuffle in the hands: Instead of riffling two packets together, sometimes the packets are interlaced on their edges via pressure. See Figure 2.4. Skillful card handlers can have the cards perfectly interleaved with this method (see Section 7.1).

Butt shuffle on the table: A tabled version of the butt shuffle is occasionally seen. We illustrate this in two variations. In each, the deck is cut into two equal piles and the piles are pushed together. In the first variation, the cards interlace, by pressure from the top down. See Figure 2.6.
2.1. ILLUSTRATIONS OF SOME SHUFFLES

In the second variation, the cards interlace from the bottom up. See Figure 2.6. With practice, this can be done perfectly and simulates a normal riffle shuffle. It is one of the most difficult feats in the sleight of hand repertoire.

It should be noted that the physics, dynamics, and probability distribution of these butt shuffles seem quite different than the models analyzed in Chapter 3. Making a believable model for these shuffles is an open problem.

2.1.2. One-handed shuffle in the hand. Magicians and card manipulators have learned to do these butt shuffles one-handed as illustrated in Figure 2.7.

Figure 2.5. Butt shuffle on the table, variation 1.

Figure 2.6. Butt shuffle on the table, variation 2.

Figure 2.7. One-handed shuffle in the hand.
2.1.3. Multideck shuffles. As will be discussed in Section 7.5, a variety of techniques is used to shuffle large decks. The following standard procedure, illustrated in Figure 2.9, may be the most popular.

With all of the cards of the multideck on the table, cut off two packets (roughly 52 cards in each) and riffle shuffle them together. Push this forward, take about 1/2 of these pushed forward cards, and cut off about 52 from what is left of the original multideck. Shuffle these two packets together and drop these on the pushed forward cards. Cut of about 52 of these pushed forward cards and about 52 from what is left of the original multideck. Shuffle these and continue until the original multideck is exhausted. Section 7.5 gives variations (and methods of analysis).

2.1.4. Overhand shuffles. This is the second most popular method of shuffling cards. The mathematics of these shuffles is developed in Chapter 13. Again, there are several variants in active use.
Standard overhand shuffle: Here, packets are dropped one onto the next as illustrated in Figure 2.10. A typical overhand shuffle of 52 cards may have 5 to 10 packets.

Strip cuts, a tabled variation: In casinos, one frequently sees the dealer strip the packets from the top of the deck into a pile on the table as shown in Figure 2.11. Observe that the dynamics are the same as for the in-the-hands variant of the one-handed shuffle (2.1.2).
Hindu shuffle: In India, a frequently used variation of the in-the-hands over-hand shuffle is illustrated in Figure 2.12.

Figure 2.12. Hindu shuffle.

2.1.5. The wash or smoosh shuffle. In casinos worldwide, and particularly in California’s legal card clubs, cards are mixed together by sliding them around on the table with both hands. After a while (usually about 30 seconds) the cards are collected together, cut, and dealt. See Figure 2.13. We present some analysis for smoosh shuffles in Chapter 14.

Figure 2.13. Smoosh shuffle.

2.1.6. Faro or perfect shuffles. Magicians and professional gamblers learned to riffle shuffle by cutting the deck exactly in half and then dropping cards one at a time, alternating exactly from each pile. Such shuffles are called Faro or perfect shuffles. There are two versions: an in-shuffle, where the original top card goes inside, and an out-shuffle, where the original top card is left outside. As explained in Section 7.1 this can be used for cheating or card tricks. It is not a good method of mixing, and as explained in Section 7.1 52 cards recycle after eight out-shuffles.

These perfect shuffles look like the Butt shuffles illustrated in Figures 2.4, 2.5, and 2.6 but the alternation is perfect. With care and practice, the cards will alternate perfectly as the shuffle is carried out.
Reverse Faro shuffles: By a reverse Faro shuffle, we mean the ordering of the cards obtained by performing a Faro shuffle and taking the inverse of the resulting permutation. So the cards in the odd positions are removed, keeping them in their same relative order, and (for out) placing them all back on top, and (for in) placing them all back under the even cards. See Figure 2.14. The point of reverse Faro shuffles is that they are easy for a nonprofessional card handler to accomplish, whereas any kind of Faro shuffle takes lots of practice.

If you try this with a packet of 8 cards (each time leaving the original card back on top), you will find that the deck recycles after 3 reverse Faros. Similarly, \(2^k\) cards recycle after \(k\) reverse Faros. See Section 7.1 for more. A sloppy version, illustrated in Figure 2.15, is a different way of actually mixing.

Figure 2.14. Reverse Faro shuffle.

Figure 2.15. Sloppy reverse Faro shuffle.
2.7. Other neat shuffles. Gamblers and magicians use a variety of other apparent mixing schemes. See Chapter 7 for a mathematical analysis.

**Monge shuffle:** This classical scheme (Monge, 1773) rearranges by passing cards from one hand to the other, putting successive cards over, then under, then over, . . ., as shown in Figure 2.16.

If randomly sized packets are pushed off instead, this mixes cards in a shuffle that has yet to be analyzed.

![Monge shuffle](image)

**Figure 2.16.** Monge shuffle.

**Milk shuffle:** Here the cards are successively milked off the top and bottom of the deck in pairs; see Figure 2.17.

![Milk shuffle](image)

**Figure 2.17.** Milk shuffle.

It turns out that this is the inverse of the Monge shuffle above.
Down and under (the Australian shuffle): Completing our trio of common neat shuffles, a packet is rearranged by successively dealing the top card down on the table, the next is put under the other cards, then down, then under, and so on until all the cards are in a single pile on the table. See Figure 2.18.

![Figure 2.18. Down-and-under shuffle (the Australian shuffle).](image)

2.1.8. Cutting the cards. Even the simple act of cutting has variations.

*Table cut:* The most frequent method is illustrated in Figure 2.19.

![Figure 2.19. Table cut.](image)

*In-the-hands cut:* This is done by amateur dealers as part of the mixing process. See Figure 2.20.

![Figure 2.20. In-the-hands cut.](image)
Multiple cut: One sees the deck cut into multiple packets, usually three. This uses a crude version of the strip cut discussed earlier. See Figure 2.21

![Figure 2.21. Multiple cut.](image)

This makes more than 20 shuffles, enough to write a book.

2.2. Some early history

Playing cards, in the Western world, date from at least 1377. Slightly before then, edicts against gambling have lists of games but no mention of playing cards. An edict of 1377 mentions cards as a recent scourge (see Chapter 4 of [378]). Shortly thereafter cards are mentioned frequently. The earlier history (perhaps through China to India to Egypt) is murky. The Wikipedia entry for playing cards has details. Let us take 1377 as a first appearance.

It seems that the early decks of cards were similar to today’s: four suits, 52 cards in all, number and court cards. Of course, this was before printing so cards were probably costly, hand-printed things which were awkward to handle much less shuffle. Playing cards were among the first things printed. Indeed, the “master of playing cards” (the first major master in the history of printmaking) produced printed cards in the 1430s; see the Wikipedia entry for “master of playing cards”.

Cards are used for gambling, for play, and for fortune-telling. It seems clear that a key aspect of all of these uses is randomization or shuffling. We set out on a quest to find some history of shuffling, particularly riffle shuffling, a main topic of our book. We are dismayed to report that the earliest clear description of riffle shuffling we can find is in 1853, close to five hundred years after cards first appeared. Here is what we found:

Cardano: The earliest mention of any kind of shuffling we could find is in Gerolamo Cardano’s (1501–1576) Liber de Ludo Aleae (the book on games of chance). This book was published posthumously in 1663, but the great Cardano scholar Ore says that Cardano wrote this book around 1520 while he was a student at the University of Padua. Cardano was an amazing character: doctor, mathematician, astrologer, and gambler. His biography “The book of my life” is fascinating. His
Liber de Ludo Aleae is the first systematic attempt to build a theory of probability. The story of Liber de Ludo Alea is well told in Ore's “Cardano the gambling scholar”. Cardano’s book uses cards and dice as examples for basic calculation. It delineates sophisticated methods of cheating and describes an early slight of hand conjuring performance in some detail. He says:

“When you suspect fraud, play for small stakes, have spectators, shuffle the cards instead of merely collecting them, and if another collects them without shuffling, he is acting fraudulently.”

That’s our first mention of shuffling, close to 150 years after 1377. Cardano’s reference to “merely collecting them” suggests the possibility that routine procedure was to haphazardly gather the cards, without shuffling, and get on to the next deal. Remember this is well before any systematic understanding of randomness. People seemed to accept the twists and turns of life as something beyond understanding (The Rumblings of the Gods Above) [203]. If probabilistic thinking didn’t exist, maybe shuffling wasn’t common. Please notice that Cardano gives no details about how the cards are shuffled.

This coupling of shuffling with cheating seems endemic. The earliest reference to shuffling in the Oxford English dictionary is to John Northbrooke (1577), “A treatise against dicing, dancing, plays and interludes with other idle pastimes.” Northbrooke writes, “They have such slights in sorting and shuffling at what game ye will, all is lost aforehand, especially if two be confederate to cousin the thirde.”

It is worth remembering that, in some card games, it is forbidden to shuffle between deals. This is the case with the popular French game Belote and with versions of bridge (Bagel bridge). The cards are to be picked up after play, gathered into a single pile, and the deck cut as usual. Then the next hand is dealt. This gives opportunities to a skillful player with a good memory.

Reginald Scot and overhand shuffles: The first clear description of a shuffling method we found comes 64 years after Cardano in Reginald Scot’s “Discoverie of witchcraft” (1584). Scot wrote a revolutionary, courageous book. At that time, “witches” were being burned and superstitions ruled. Scot was a rationalist who argued that just because someone acts strangely or seems impaired doesn’t mean they are in league with the devil. He had to be very careful about what he said regarding religion. Indeed, his book was ordered burned by James I (of the King James Bible).

Scot studied all kinds of magic, including conjuring. He took lessons and wrote an expose which is the first expository account of natural magic (that is, magic which deals with natural forces directly, as opposed to ceremonial magic which deals with the summoning of spirits). Among the 52 tricks Scot describes (such as transforming or altering the color of one’s cap or hat), there is a sophisticated description of card tricks. This begins (Scot, page 332):

“But in shewing feats, and juggling with cards, the principall point consisteth in shuffling nimblie, and alwaies keeping one certaine card either in the bottom, or in some knowne place of the stacke, foure or five cards from it.”
He then goes on to give a page of full technical details of manipulation with
the standard overhand shuffle, replete with in jogs, out jogs, and other niceties.
This suggests that the overhand shuffle was in common use in 1584. (If, during an
overhand shuffle, a card is shuffled “out of line” with the others, so it can be kept
track of by a nimble fingered cheat, the card is said to be in (or out) jogged.)

According to Hoyle? One natural place to look for shuffling details is in
books of rules for card games. Surely a book such as Hoyle’s “A short treatise on
the game of whist” (1742) would talk about shuffling. Alas, none of the books of
Edmond Hoyle (1672–1769) contain any kind of rules or shuffling instructions. They
are rather tips and tricks for efficient play. There are many other “rules for card
game” books and a magnificent bibliography: Manfred Zollinger’s Bibliographie der
Spielbücher des 15. bis 18. Jahrhunderts (Hiersemann, 1996). We have examined
many of these. The results are disappointing. For example, one of the earliest
English rule books, Charles Cotton’s “The Compleat Gamester” (1674), sometimes
mentions shuffling as part of a game’s description. These mentions are never more
than (we quote):

“The dealer shuffles and the other cuts.”

That’s it! Many editions later, a variant by Richard Seymour (“The Court
Gamester” (1719)) adds some detail:

“The dealer is to shuffle the cards and offer them to the other
to cut. If in cutting, he should scatter or anyways displace the
cards, they are to be shuffled and cut again. If the person who
does not deal has a mind to shuffle the cards, he may; but the
dealer is to give them the last shuffle.”

This quotation shows that there was some “mixing sophistication”. Here is
perhaps our most hopeful find: a series of French rule books was published begin-
ing with La Mariniere, J. (1654) La Maison Academique. In a later edition we
find:

“Lors celui qui a coupé la moindre carte mêle les trente-deux, +
les presente à l autre qui en fait deux parts, pourvu que ce soit
nettement, car qui couperoit en esparrillant les cartes, cela ne
vaut rien, + sauf remêler + donner à couper, pour en faire deux
parts, les quelles etant rassemblées par celui qui les a presentées,
....”

The translation of this paragraph is essentially word for word a translation from
Seymour just above. The old books all copied from each other.

Cheat shuffling: Early books on cheating at cards describe systematic ways
of “mixing” that have the appearance of shuffling but are actually ways of cheating.
For example, in the anonymously authored “The whole art and mystery of modern
gaming fully exposed and detected containing an historical account of the secret
abuses practiced in the games of chance” (1726), we find cards carefully mixed
“upon drawing them through my hands from top to bottom” (milk shuffle), followed
by “now shuffle and put every card over and under” (Monge shuffle). It is not hard
to see that doing a milk shuffle followed by a Monge shuffle leaves the cards in order.
Thus a sophisticated full deck false shuffle is being described. This is followed by the intriguing:

“You are to consider that the banker commonly shuffles and cuts the cards at Bisset, and in order to prove what I have often said, viz. that a pack of cards may be changed into any form whatsoever, and that even by a school-boy when the first principles are set down, I will give the following example.

Suppose \(a, b, c, d, \ldots, e, f, g, h, \ldots\) to be certain cards best known to yourself. It can not be thought a difficult task to join \(a\) to \(e\), \(b\) to \(f\), \(c\) to \(g\) and \(d\) to \(k\) and further to continue to 52 in this same order” (page 91).

Note that in the above quotation, Bisset is the name of an old French card game.

We wonder if this is not an early description of a Faro shuffle (perfect riffle shuffle). After all, if an 8-card deck is set up from top to bottom as \(a, b, c, d, e, f, g, h\), the cards are cut into two piles \(a, b, c, d\) and \(e, f, g, h\) and then perfectly interlaced, they wind up \(a, e, b, f, c, g, d, h\). The milk, Monge, and perfect shuffle are all intimately related (see [161] for definitions and details). There is some dispute about the above as a first description of perfect shuffles in the magic literature (see Denis Behr’s webpage on conjuring credits, under Faro shuffles). We have given the details and the reader can decide. The discussion above was reprinted, word for word, in Richard Seymour’s Complete Gamester (1734) in many editions.

**Riffle shuffling:** The first clear description of riffle shuffling that we have found appears in the literature on magic tricks. J. W. Ponsin, a serious amateur conjuror, wrote “Nouvelle Magie Blanche Devoilee” (roughly “new white magic”) in 1853. This is the first serious, popular, modern magic book. It was copied, extended, and plagiarized in many languages, most notably in English, as Professor Hoffman’s Modern Magic (1876).

In his section on card tricks, Ponsin (1853, page 39) describes “false shuffles”. As a preamble he says:

“You will see, if you pay attention to it, that good card players, those who are in the habit of handling cards, mix them by making two packets of the deck interlace, the one into the other, which they then repeat several times and which they execute by moving the fingers to push the cards down to interlace them into the other packet. Then, with a blow of the hand, the fingers on the upper end of the pack and the thumb on the lower end, they square the cards.”

The paragraph above follows a translation by Jean Hugard, a modern dean of magicians. In a footnote, Hugard adds: “Note that the riffle shuffle was unknown at this date”. Hugard’s translation is published in the conjuring periodical Gibeciere (Vol. 15, No. 1, 2020, pages 81–168).

A second early description of riffle shuffling appears in the literature of cheating at cards. Jean-Eugene Robert-Houdin, the great French magician (Houdini “borrowed” his name) wrote Les Tricheries des Grecs (Paris, 1861). His Chapter XIII on false shuffles carefully explains how to cheat during riffle shuffles. This suggests riffle shuffling as a commonly used method.
Conclusions: The bottom line is that there is no careful history. People just didn’t care to record this detail. It is similar with other methods of randomization. For example, the age-old fortune-telling ritual called the I-Ching (book of changes) depends on generating a random binary six-tuple. There are many ways of doing this, by flipping coins, using a bundle of sticks, and drawing cards from a fortune-telling pack. Users treat these indifferently, even though they give markedly different distributions for outcomes; see Chapter 8 of [158] for extensive discussion.

Here is a final illustration for “nobody cares about shuffling”. Perhaps the best history of playing cards is Michael Dummett’s magisterial “The Game of Tarots” (Duckworth, 1980). Dummett (1925–2011) was a celebrated Oxford logician. During a “down time” he became fascinated by a legendary story about the gypsy invention of Tarot cards. The story goes that at a particularly rough time, the gypsies, fearing being wiped out, put their lore, wisdom, and images into Tarot cards. Dummett shows convincingly that this story is all made up. The story first appears in 1826. Tarot cards have been used for games from 1470 to today. Along the way, Dummett compiled a fantastic detailed history, the best we have, of all types of playing cards. Nevertheless, a careful perusal reveals not a single mention of when or how to shuffle. Dummett also wrote a popular book trying to keep Tarot games alive: “Twelve Tarot Games” (Duckworth, 1980). Each game has a perfunctory, “After the cards are shuffled, the dealer gives them to the player on the left to be cut.”

Above we have reviewed the history of shuffling playing cards. There are a dozen related historical questions regarding the history of games, randomness in games (e.g., chess used to have its opening moves randomized and, just recently, as checkers fades into oblivion, adherents tried to save the game which most often ended in a draw among skillful players by starting with random openings), playing cards, individual games, fortune-telling with cards, Tarot cards, ... The Wikipedia entries on these subjects provide a good start as do the wonderful publications of Thierry Depaulis and the journal “The playing-card”.

The website “The world of playing cards” provides endless further links as well as links to the amazing world of people who collect old playing cards and the lucrative world of people who design new playing cards. Of course, other games (dominoes, Mahjong) need shuffling too.

The history of shuffling just presented and the analytical results later in the book lead to a fascinating conclusion. Riffle shuffling is a late arrival (1853), despite being the only method of shuffling that leads to a truly mixed deck in a reasonable number of iterations. In Chapter 8 we show that roughly seven riffle shuffles suffice for mixing. Before this, the prevalent (only?) available method of mixing was the overhand shuffle or its variant, the over-under shuffle. In Chapter 13 we show that this takes roughly ten thousand repetitions to mix. All of this means that for roughly the first five hundred years that playing cards were used, people played with badly nonmixed cards.
Applications to magic tricks, traffic merging, and statistics

The basic structure of riffle shuffles has been used for centuries by crooked gamblers and for at least 100 years by magicians. This chapter describes some magical and gambling uses. We are not trying to teach card tricks (although some good ones occur). For this, see 158.

The chapter is organized as follows. The first ten sections provide applications to magic and gambling. Of these, the first eight sections focus on riffle shuffles. Sections 16.9 and 16.10 treat the overhand and smoosh shuffles, respectively. The last two sections of this chapter are somewhat speculative; Section 16.11 describes an application of shuffling to traffic merging and Section 16.12 describes an application of riffle shuffles in statistics.

16.1. Rising sequences

Rising sequences are a basic invariant of riffle shuffles that was discovered by the magicians C. O. Williams and Charles Jordan early in the 20th century. If a deck of cards ordered 1, 2, ..., n from the top down is given a single riffle shuffle with k cards cut off, then cards 1, 2, ..., k and cards k + 1, ..., n are in increasing order, thus forming two rising sequences. After two shuffles, there are (at most) four rising sequences, and so on.

Williams 478 used this to do a “card reading”. A prearranged deck was riffle shuffled once by the spectator. The deck had its back design all aligned in the same direction and the handling forced the top cards “turned end to end” so the two rising sequences could be identified from the back.

Jordan took things to another level, advertising “Mindreading by Mail” in 1916. The performer sends the spectator a deck of cards with the following instructions: “Take the deck out of the case. Give it a simple cut. Give it a riffle shuffle. Give it another cut and another shuffle. Give it a few more cuts. I’m sure you agree, nobody can know the name of the top card. Please look at it and remember it (say, it is the six of hearts). Take the top card and put it into the middle of the deck. Give the deck another riffle shuffle, and then give it a few more cuts. Please mail the deck back to me; oh yes, at six o’clock every evening, please concentrate on your card.” A few days later, the spectator gets a letter back saying, “It was the six of hearts.”

How does the trick work? To understand consider a simplified version of the trick with 13 cards, initially arranged

A 2 3 4 5 6 7 8 9 10 J Q K.
Consider a “one-shuffle version”. The 13-card packet is repeatedly cut, the top card noted and inserted into the center of the packet. The packet is freely cut again and a single riffle shuffle is performed followed by a few more cuts. Suppose the cards are now in order

10 9 4 J Q 5 6 K 7 8 A 2 3.

Can the reader look at this arrangement and tell what card was selected? (Please try it!)

The key is to start at the original top card (the ace) and sequentially follow along (going “around the corner”) counting

A 2 3 (skip the 10, 9) 4 5 6 7 8 (around the corner) 10 stop.

The 9 is out of order. It should be between the 8 and 10. So, the selection is the 9.

Practicing this can make a kind of solitaire! We invite the reader to determine the selection from

10 J Q K 7 A 8 2 9 3 4 6 5

or

A 2 3 4 9 J 10 Q K 5 6 7 8.

It also works after two shuffles. If you like this kind of solitaire, you can check if you got the right answer by using cards with one way backs, performing the cuts and shuffles with the cards face down but turn the selection end for end (so it has opposite orientation). Then study the cards face up and check your answer using the orientation.

All of this explains the idea: after a riffle shuffle the deck is in two “chains” and the procedure forces the selection into a third chain of length one. Turn now to the full “mindreading by mail”.

The deck is prearranged in a known order (perhaps by just recording the order of a shuffled deck). For now, say, it is in order 1, 2, ..., n. After three riffle shuffles, the deck has eight rising sequences. When the spectator removes the top card and inserts it in the middle, it makes a ninth rising sequence of length one. So, when the performer gets the deck back, here is what to do. Start turning the cards up, one by one, as if playing solitaire. Say, the top card is the eight of diamonds. If the next card is the nine of diamonds (next card in the arrangement), play it on the eight. If not, start a new pile. Continue. If you turn up a card adjacent to a card that is showing, play it on that card. Otherwise start a new pile. When you finish, you will find eight piles, each containing about an eighth of the deck, and a ninth pile, of size one, containing the selected card.

The cuts don’t really mix cards, they just leave things in a cycle (so in reckoning the next card you may have to go around the corner). The trick as described is not “sure fire”. For example, a difficult spectator may put the original card back on top (or very near the top or bottom), not making a new rising sequence. Practical performance details (and the statistics of success, nearly 85 percent, rising to 95 percent if only two shuffles are used) are in [46].

With all of the shuffling knowledge we have acquired, we can add a few original twists to Jordan’s trick. Consider a one-shuffle version. The deck is repeatedly cut and the performer instructs, “Cut into two piles, about equal.” Could I know the name of either of the top two cards? (No.) Two spectators each remember one of each; the top cards are inserted (randomly) in the middle of the opposite halves. The two halves are now riffle shuffled together, and the deck is then cut randomly
a few times. The performer looks through the deck and determines both selections. 

If the reader looks at Section 7.5 above (shuffling big decks), they will see that the 
deck can be cut into four piles, the top card of each noted. These top cards are 
each inserted into a different pile. Shuffle two of the decks together, the other two 
piles together, and finally riffle shuffle the two half-decks together. The spectator 
gives the deck a final random cut. All four cards selected can be determined. 

Jordan gave a host of other uses for rising sequences which may be found in 
the “Shuffle systems” chapter of the book about his best card tricks. Jordan 
advertised his trick in the May 1916 issue of the magic magazine *The Sphinx* for 
$2.50 (for 50 cents he would perform it by mail, and you could buy the secret for 
$2.00). At the time, an issue of *The Sphinx* sold for 10 cents. It was an expensive 
trick. 

Cruder uses of rising sequences are easy to exploit. Take the 13 spades and put 
them in the following order at the bottom of the deck: 

7, K, 2, Q, 6, J, 3, 10, 5, 9, A, 8, 4 

(so the 4 is at the bottom of the deck). Have a spectator riffle shuffle twice and then 
inspect him to remove the spades. The order is preserved with high probability. 
Have the spectator pick up the spades (keeping them face down) and do an “under-
and-down deal”: put the top card under the packet, deal the next on the table, put 
the current top card under the packet, deal the next onto the card already on the 
table. Continue under, down, under, ..., until all cards are on the table. You will 
find the spades are now in order. 

Alternatively, separate the four suits into four piles, each suit arranged in some 
clever order. Have the spectator riffle shuffle the spades and hearts together (two 
13-card packets). Have him shuffle the clubs in (a 26-card packet with a 13-card 
packet) and finally shuffle the diamonds in. To the uninitiated, the cards have been 
shuffled three times, but each suit is in an unchanged relative order. Jordan himself 
arranged


After the shuffle he would have the spectator name any suit and remove all cards of 
that suit (keeping them in order). With the 13-card packet face down, the spectator 
spells A-C-E, transferring a card from the top to the bottom for each letter. The 
spectator turns up the next card and shows the ace. The spectator then spells 
deuce, then three, and so on. Each time the next card is of the proper value. 

16.2. The Gilbreath principle 

Get a deck of cards and arrange it so that the colors alternate red, black, red, 
black,.... Put the deck face up on the table, and cut it into two piles so opposite 
colors show. Riffle shuffle these two piles together. Turn the deck face down. Turn 
up the two top cards. One will be red, one black, similarly for the next two, and 
the next two, right down to the bottom pair. This is called the first Gilbreath 
principle. 

Here is a variant, called the second Gilbreath principle. Remove the four aces, 
four twos, threes, fours, and fives (20 cards in all) and arrange them (from top 
down)

1, 2, 3, 4, 5, 1, 2, 3, 4, 5, 1, 2, 3, 4, 5, 1, 2, 3, 4, 5.
Holding the 20-card packet face down, deal a pile of any number face down onto the table, reversing the order as you deal. Riffle shuffle the pile on the table with the pile left in the hand. You will find (no matter how you shuffle) that the top 5 cards are a 1, 2, 3, 4, 5 (in some order). The next 5 are a \{1, 2, 3, 4, 5\}, same for the last two sets of five.

A mathematician may want to understand these principles and could ask, “What’s going on? Just what kinds of structures are preserved?” A complete answer and many performance tricks, as well as a connection with the “Mandelbrot set” of dynamical systems, can be found in Chapter 5 of [158].

16.3. Tops and bottoms are special

A feature of riffle shuffling is that the top card tends to stay near the top. Of course, this can be exploited by the shuffler who “just happens” to finish a riffle shuffle by dropping the last few cards from the original top of the deck (or starting the shuffle so the original bottom cards drop first, thus leaving them on the bottom). Less obviously, suppose the two halves of the deck, before the shuffle, each have five spades on top. After any kind of shuffle, the top 5 cards are sure to be five spades (a flush in poker). If each half starts with four of a kind on top, after a single riffle shuffle there will be a good poker hand on top (at least a full house). To see how these ideas, along with the Gilbreath principle, can be exploited, see pages 66–67 of [158].

16.4. A performable magic trick

We now combine the ideas above to create a good, performable trick. The trick will reveal a novel regularity of shuffling, not recorded above.

The trick requires that the spectator knows that in poker you get a hand of five cards and that hands are values as follows:

- no pair (then high card wins),
- one pair (then higher pair wins),
- two pair,
- three of a kind,
- straight (e.g., 5 – 6 – 7 – 8 – 9, not all of same suit),
- flush (all five of the same suit),
- full house (three of a kind and a pair),
- four of a kind,
- straight flush (five cards in a sequence, all in the same suit),
- royal flush (e.g., 10-J-Q-K-A all of the same suit).

To start the trick, the performer asks a spectator if they would like a lesson in how to cheat at cards. When someone agrees, let’s call her Gael, the performer takes a deck from the card case and says, “It’s all about skillful shuffling and dealing. Do you know how to play poker?” She says, “Yes.”

“Gael, we are playing 5 handed, so please deal the top 25 cards off the deck onto a pile on the table. Good, now pick up the dealt cards and deal a pile of about half onto the table. Can you riffle shuffle? Good, please riffle shuffle the two packets (cards in the hand and cards dealt off) together. OK, now, just for practice, deal 5 hands, one card at a time dealing around, giving yourself the 5th hand. The aim is to give yourself a good hand and also to give someone else a good hand (slightly less than yours). Let’s look.”
Turn the hands up one at a time (keeping them in the same order). Probably Gael has a poor hand, and someone beats her. Turn the hands all face down. Have her put the packets one on another to have all 25 cards in a packet (the way this is done doesn’t matter). Finally, have her put the top card onto the bottom of the deck. Continue: “OK, that’s practice; now for the real deal, try to give the fourth player a good hand, but make sure you beat them!” Gael deals 5 hands again. When the fourth hand is turned up, it’s a straight, a very good hand. But Gael has dealt herself a straight flush.

That’s the bare bones of what the trick looks like. How does it work? The deck is set up beforehand. The top 25 cards are in prearranged order with 5 hearts on top, 5 hearts on bottom, and three straights, in order, in between (suits don’t matter for the straights). Just to be definite, say they are (from top down)

\[8H - 7H - 6H - 5H - 4H - A - 2 - 3 - 4 - 5 - A2 - 3 - 4 - 5 - A - 2 - 3 - 4 - 5 - KH - QH - JH - 10H - 9H\]

OK, the trick now works itself (!). These 25 cards are on top of the rest, and all are in the card case. Take the deck out and have the spectator deal the top 25 cards off and proceed as in the description above. The only thing to be careful about is that when you say, “Deal about half the deck onto the table”, they must deal at least 10 and at most 15 cards off. Try it.

The new shuffling fact alluded to above is that two sets of hearts will combine together to form a straight flush (not just a flush). We will let the reader see this for themselves. We learned this fact from the mathematician-magician Matt Baker.

A reader with a magic background will have ideas about adding some sleight of hand and other subterfuges to improve the trick. Our original version is in Chapter 5 of [158]. Matt Baker’s version is in his book [41] (the trick in question is called “Gilbreath poker”). A splendid sleight of hand version is on pages 7–8 of [49].

16.5. A homework problem

Just to see if you’ve got the ideas and to give you an example of how they are used, let us describe a very good magic trick along with how it works. We ask if you can understand why it works.

The effect: The performer removes two cards from the deck, writes something on each of them, and puts them back. A spectator gives the deck a riffle shuffle. The spectator then removes cards two at a time (first the top pair and then the following pair, and so on), turning them up. Matching red pairs go in one pile, matching black pairs go in a second, and mixed red-black pairs go in a third. It’s like in basketball “two points for the red team”. When the pair containing one of the prediction cards is turned up, the spectator reads aloud:

“You have just dealt the fourth and last pair of matching black cards. However—”

Continuing until the second prediction:

“— you are destined to deal two more matching pairs, and both will be red.”

Pairs are dealt until the deck is finished. Both predictions prove correct.
**How it works:** The deck is arranged: the top 26 are \(RB\)–\(RB\)–\(\cdots\)–\(RB\) as alternating in pairs. The bottom 26 are (with 2C, 2S the two of clubs and spades)

\[
BBBB - BR - BB - BR - RR - RR - B - 2S - BRBR - 2C - RRRRR.
\]

To perform, remove the 2S from the deck and write the first prediction on its face. Replace it in its original position and do the same with the 2C (for the second prediction). Cut the packet at 26 and have the spectator shuffle. The rest works as explained.

This trick first appeared in the magic journal *The Pallbearer’s Review*, Volume 2, Number 9, July 1967 (page 111). If you can explain why it works, you will have understood some things about riffle shuffles.

### 16.6. Riffle stacking

Magicians and crooked gamblers have developed a variety of techniques for “arranging” cards during a riffle shuffle. This appears in the first published description of riffle shuffling (see Section 2.2 for some of the history) and continues to this day.

The following description assumes that the cards are shuffled “on the table” as pictured in Figure 2.3 in Section 2.1. It also works for the usual way of riffle shuffling in the hands.

The idea is simple. Suppose four aces are on top of the deck with the deck face down on the table. Cut the top half of the deck to the right and commence a riffle shuffle. As you get to the top, hold the top 4 cards with the left thumb and the top 3 cards with the right thumb. Drop the left 4 under the right 3 and complete the shuffle. This results in 3 aces on top followed by 4 cards followed followed by the fourth ace. Refer to this as “shuffle 4 under 3”. Continue to “shuffle 4 under 2”, then “4 under 1” and finally “4 on top”. These four shuffles stack the four aces every fifth card: if five hands are dealt, the dealer gets the aces.

To do this skillfully, so that the keenest observer will not suspect, let alone detect, the deception, requires years of practice. There are hundreds of practitioners who have put in the practice, and we have friends who can do it blindfolded. To stack for \(k\) hands change 4 to \(k - 1\). Of course if \(k = 8\), this can be done in two stages (to start “4 under 3” and then “3 under 3”).

There is not much mathematics above but you can make math anywhere. Alex Elmsley worked out a riffle stacking sequence in which the same shuffle was simply repeated (instead of “4 under 3” and then “4 under 2”, \(\ldots\)). This was independently discovered by Steve Freeman. Here is how it goes. Begin with 4 aces on top. We copy Elmsley’s description on page 393 of [352]. The tabled cards are cut into two with the aces on top of the right-hand packet. Shuffle until the right and left each have 4 cards left. Then, drop 1 from the right, 3 from the left, 3 from the right, last from left on top. Push the packets together. Repeat this exact sequence twice more (three shuffles in all). You will find the aces come to the dealer if four hands are dealt.

The above only scratches the surface of what can be done with skillful riffle shuffles. The diligent student will find a post graduate course in Edward Marlo’s privately published series “Riffle shuffle systems” (Chicago, 1959), “The patented shuffle” (Chicago 1964), and “Riffle shuffle finale” (Chicago, 1967).
16.7. Close stays close

In a riffle shuffle or two, cards that are originally together stay close together. This may be exploited in various ways. Gamblers can use the expectation that the following card will be appearing shortly. Our formulas for features in Chapter 4 give precise probabilistic descriptions. For magical use? Well, clearly if you spot the bottom card of the deck, have a card selected and replaced on top, and the deck is cut a few times, then the noted card is just above the selected card. Now if the deck is riffle shuffled once or twice, the selected card is just a few cards below the noted card. The magician must find a way to end, perhaps by asking a judicious question or two (let’s see, I think it was a red card . . .). For those reading with a magic background (and access to a library) more sophisticated uses can be found in three tricks by Herb Hood in *Jinx*, # 136 (April, 1941). Better still, use your imagination.

16.8. Reds and blacks

A different invariant of the riffle shuffle can be understood by taking a 52-card deck and arranging it so that all the red cards are together on top. With the deck on the table, cut off approximately the top 3/4 of the deck and place it adjacent to the bottom 1/4. Next cut off approximately 13 cards from the large packet and place these cards on top of the original bottom 1/4. This leaves you with two packets of approximately equal size. Riffle shuffle these two together (or have a spectator do it).

If you look, you will find that almost all of the top half is red and that almost all of the bottom half is black. And any intermixing occurs near the middle. This may be exploited by having two spectators each take cards: one from “near the top”, one from “near the bottom”. Have them replace their cards, freely, but in opposite halves. Looking through the deck, it is easy to spot the two selections. Further, the shuffle above may be repeated after the selections are returned with small chance of disturbing the final revelation. We don’t know the first appearance of this. A relevant reference is page 54 of “Dai Vernon’s more inner secrets of card magic” by Lewis Ganson. A more subtle application is “Galaxy” by Wyman Jones and Paul Harris, published in Book 3 of Harris’s “The Art of Astonishment”.

16.9. Overhand shuffle

As explained in Chapter 13, the overhand shuffle, though widely used, is a poor way to mix up cards. This fact has been exploited centuries ago as a way of cheating. In the context of magic we make two observations. First, after an overhand shuffle, cards near the top wind up near the bottom after one overhand shuffle and back near the top after two overhand shuffles. Second, the overhand shuffle tends to keep pairs together. Thus if a selected card is placed adjacent to a known card, an overhand shuffle or two, even by the spectators, will usually leave this unchanged. Then, running through the deck, it is easy for the performer to find the selection. In Chapter 13 we detailed ways that the overhand shuffle can be used to “stack” the cards for a magician’s demonstration of a gambling game. For those who can find it, we recommend Dai Vernon’s “Lucky pair” in [138].

A nice overhand shuffle version of the red/black riffle shuffle of Section 16.8 was devised by Charles Jordan in his booklet “Thirty card mysteries”, privately
published in 1919. Arrange the cards so that all the red cards are together on the top. Do a normal overhand shuffle but, as you get towards the middle, start running the cards singly, simply reversing their order. When you are through the middle, overhand shuffle normally. You will find all the red cards together at the bottom. This can be used for the tricks suggested in Section 16.8.

We are sure that there is more structure to discover for overhand shuffles. We encourage you to explore further.

### 16.10. Smooshing

A wonderful example of “smooshing as magic” occurred in the late 1970s. A talented young “psychic” Susie Cottrell appeared as a guest on one of America’s most popular TV shows *The Late Show with Johnny Carson*. Susie fooled Carson (a former professional magician), thousands of other magicians, and millions of the American public. An ordinary deck was spread around the table. A spectator selected 5 cards while the spreading occurred. Susie backed away and said, “Concentrate on the lowest of your cards”, and named the card. This was repeated in many variations in a 17-minute slot, quite a long time.

Susie made the rounds of psychic investigators, finally running into arch-skeptics Martin Gardner and James Randi. Their report appears in the Spring 1974 issue of the *Skeptical Inquirer* (with followups in Spring and Summer 1980 and Spring 1981). The following details are instructive: “...she suddenly banged the deck on the table and began spreading the cards wildly about on the surface of the table with both hands in a circular fashion, rearranging them repeatedly and ‘patting’ them into seemingly random patterns. She pushed them into various configurations while she spoke, asking her subject, a young woman she’d chosen to sit opposite her at the table, to reach into the spread and select any five cards. As each card was slipped out of the spread, Susie would rearrange them somewhat, seemingly at random.”

This time, 4 of the cards were somehow eliminated and the last shown to match a prediction Susie had previously written. The investigators explain the method: they saw Susie peek at the top card and write it down as her prediction. Then:

“We noted, the top card almost immediately got tossed off to one side, close to Cottrell and near her arm, which would cover it during the fussing about. It would be picked up by one thumb, slid about a bit, then dropped off again in a different position, but always in full control and available.

“Finally, Cottrell introduced the chosen card into the main mass of cards, arranging it all by patting and pushing so that the one important card lay in a position where it was most likely to be chosen....”

Independent reports by our friend, card expert Steve Freeman, substantiates this account. It’s not math, but it shows how smooshing blends into magic.

Smooshing cards around makes it easy to cheat and manipulate cards — leaving a few cards “off to side” while others are mixed and so on. We have not seen organized efforts to stack cards while smooshing (as explained above for riffle or overhand shuffles). A celebrated development in magic circles is called “the Schulien force”. Magic ethics prevent further description, but those who can may consult “The magic of Matt Schulien” by P. Willmarth [479].
16.11. An application of shuffling to cars merging in traffic

This is a speculative section, but we are confident there is nice mathematics hidden here. Consider two lanes of traffic merging into one. Cars usually alternate (called zipper merging in the traffic business). But sometimes, two or more proceed from the same lane. It doesn’t take much imagination to think of this as a riffle shuffle. Similarly, it often happens that a single-lane road opens out to two lanes and a single stream of cars splits or “unshuffles”. Finally, there are many long stretches of road where this unshuffling and shuffling alternate, repeating at traffic lights every mile or so.

Thus, consider \( n \) cars, in initial order 1, 2, \ldots, \( n \) in a single lane. Suppose they are unshuffled and shuffled several times, with some randomness thrown in. What happens to the cars as time goes on? If car \( i \) is \( l \) ahead of car \( j \) at the start, how long will it take to switch their positions? Of course, with real cars, all sorts of other considerations enter. Cars enter and leave the road, move at different speeds and change lanes when they can. Slower cars stay on the right. Many such bells and whistles can be added to our simple model. For now, let’s start with the following set of assumptions:

- In splitting from a single lane, the first car picks up or down with probability \( 1/2 \).
- The following cars alternate with probability \( p \) (\( 0 < p < 1 \)) and follow the previous car with probability \( 1 - p \).
- With two lanes merging into one, a random lane is chosen (probability \( 1/2 \) for each lane), and the top car in that lane “bolts ahead”.
- Successive cars merge with probability \( p' \) (\( 0 < p' < 1 \)) and wait with probability \( 1 - p' \).
- All choices are made independently.

With these ingredients settled, we have a well-defined Markov chain on the induced permutations of 1, 2, \ldots, \( n \) when the cars are in a single lane.

The following proposition is a first observation, and we thank Richard Stanley for help with its proof. For its statement, recall that a permutation \( \pi \) is called 321 avoiding if for any \( i \) less than \( j \) less than \( k \), one can’t have \( \pi(i) \) greater than \( \pi(j) \) greater than \( \pi(k) \).

**Proposition 16.11.1.** A permutation \( \pi \) can arise from an inverse riffle shuffle followed by a riffle shuffle (or equivalently from a single split/merge step) if and only if it is 321 avoiding. The number of such \( \pi \) is the Catalan number \( \binom{2n}{n}/(n+1) \).

**Proof.** First consider a deck of three cards in initial order 123. By direct enumeration of cases, after a single split/merge shuffle all orders are possible except 321. It follows that a split/merge shuffle of an \( n \)-card deck is 321 avoiding.

Conversely, let \( \pi \) be any 321 avoiding permutation. Define a partial order on 1, 2, \ldots, \( n \) with \( i \leq j \) if and only if \( i \leq j \) and \( \pi(i) \leq \pi(j) \). It is easy to verify that this is a partial order. Recall that a chain in a finite partially ordered set is a subset where all elements are comparable and that an antichain is a subset where no two elements are comparable. A classical theorem of Dilworth (Exercise 3.77 of [438]) says that in a finite partially ordered set, the size of the largest antichain equals the minimal number of chains that cover the set (as a disjoint partition). Now except for the identity, 321 avoiding implies that the size of the largest antichain is 2. Thus Dilworth’s theorem shows that there is a covering of \( \{1,\ldots,n\} \) by two chains. So
there are $i(1) < i(2) < \cdots < i(k)$ and $j(1) < j(2) < \cdots < j(n-k)$ with

$$\pi(i(1)) < \pi(i(2)) < \cdots < \pi(i(k)) \quad \text{and} \quad \pi(j(1)) < \pi(j(2)) < \cdots < \pi(j(n-k)).$$

This shows that $\pi$ is the result of a split/merge shuffle.

A theorem of Knuth shows that the number of 321 avoiding permutations in $S_n$ is the $n$th Catalan number. For this and other interesting results on pattern avoidance, see Chapter 4 of [72].

Practically, $p, p'$ will be close to 1. We note that if $p = p' = 1/2$, there is some fascinating previous work on these split/merge shuffles. Consider a deck of $n$ cards, in order from the top 1, 2, \ldots, $n$. Perform an inverse GSR shuffle; that is, flip a fair coin for each card and take all cards labeled 0 in one hand (keeping them in the same relative order) and all cards labeled 1 in the other (again keeping them in the same relative order). Now GSR these two piles together as usual. This split/merge version has been studied in unpublished work of J. C. Uyemura-Reyes. He was able to identify some of the eigenvalues, the ones appearing in low-dimensional representations of $S_n$. Alas, even for a 4-card deck, it appears that some of the eigenvalues are irrational instead of the nice positive integers of random to random (thanks to Nadia Lafreniere).

A much simpler version of this shuffle, namely random to top followed by top to random, resulting in “random to random” has led to fascinating research and to a complete analysis (see Section 8.9). We have every reason to hope that this riffle shuffle extension of random to random will be equally interesting. Maybe it will eventually say something about cars merging.


Data is sometimes collected in the form of rankings. For example, a group of people is asked to rank 5 different kinds of chocolate chip cookies, each ranker indicating favorite, next favorite, \ldots, least favorite.

With $N$ rankers, this gives rise to $\pi_1, \pi_2, \ldots, \pi_N$ in $S_5$. Various political elections involve rankings (e.g., in Australia, Ireland, San Francisco, \ldots) so $N$ can be large. For example, an Irish election studied below involved about 70,000 people ranking 14 candidates [243]. Search engines generate rankings and even in the discipline of card shuffling one encounters data consisting of many arrangements of a deck of cards.

A host of specialized techniques for analyzing ranking data is available. We have written on this subject [142], [143]. A good summary of the literature is in the book [341].

A curious entry in this literature due to Huang and Guestrin [272] uses riffle shuffles.

Their idea can be explained this way. Suppose there are $p$ items to be ranked and that each comes with a binary covariate. For example, of the 5 brands of chocolate cookies, 3 may be package brands and 2 may be from local bakeries (or Republican/Democrat for elections). It might be that voters make up their final ranking in 3 stages:

- Rank Type 1 items among themselves.
- Rank Type 2 items among themselves.
- Choose an interleaving (shuffling) pattern $\epsilon_1, \epsilon_2, \ldots, \epsilon_p$ and put the Type 1 items in ranked order where $\epsilon_i = 1$ and similarly for the Type 2 items.
For our cookie example, if three national brands (Type 1) were ranked
\[ \text{Nabisco} > \text{Chips Ahoy} > \text{Mothers} \]
and the top two bakery brands (Type 2) were ranked
\[ \text{Bakery B} > \text{Bakery A} \]
and bakery cookies were preferred to national brands (interleaving 2, 2, 1, 1, 1), the final ranking would be
\[ \text{Bakery B} > \text{Bakery A} > \text{Nabisco} > \text{Chips Ahoy} > \text{Mothers}. \]

If the 3 choices above are made independently, Huang and Guestrin say the data generating mechanism is a riffle independence model. Given data \( \pi_1, \ldots, \pi_N \) in \( S_p \), one can try to estimate the 3 probability distributions involved (and even test for independence).

If there are \( a \) types of covariates, then the GSR \( a \)-shuffles of Section 3.1 are relevant. If covariates are available, independence offers a reasonable simplifying assumption to estimating a distribution in the \( p! \)-dimensional simplex. The paper [272] develops two further themes. First, in estimating the two marginal ranking distributions above, one may use the ideas recursively in each group (if nuanced covariates are available, like light or dark chocolate).

The second of these is more speculative. We have assumed above that covariates are available for each ranked item. Suppose they weren’t. One could posit various clusters of ranked items searching for simple descriptions of the data. This becomes computationally involved and [272] develops algorithms and complexity bounds and applies their methods to simulated and real data.

If we had to analyze permutation data given with (say) binary covariates, we could try representing that data as signed permutations. So 31254 indicates that items 1, 2 and 4 are Type 1 and items 3, 5 are Type 2. Fourier analysis on the hyperoctahedral group as in [143] offers a route to data analysis.