Signaling the Key Card in Fitch Cheney: Variations on a Theme

Neil J. Calkin[1], *Eliza Gallagher*[1], *and Ben Gobler*[1]

[1] *Clemson University, Clemson, South Carolina 29634*

Abstract

We discuss various methods for signaling the position of a key card in the classic Fitch Cheney trick. We analyze the relative frequency with which key card positions arise when using various signaling conventions for a randomly selected hand of four cards, with an eye toward balancing a desire for uniformity with ease of calculation.

1 Introduction

In the classic Fitch Cheney card trick [Lee51], an unsuspecting audience member selects five cards from a standard 52-card deck. The magician's assistant studies the five cards before selecting one to hide. The assistant then arranges the remaining four cards in a particular order so that the magician, on viewing those four cards, can determine the name of the hidden card, declaring it aloud to the audience's wonder. The history of this trick, together with a thorough and delightful explanation of how it is performed, can be found in Chapter K♣ of *Mathematical Card Magic: Fifty-Two New Effects* by Colm Mulcahy [Mul13]. This trick has gained popularity in recent years after being reintroduced by Art Benjamin [Mul13, p. 312], and there have been many lovely discussions of the mathematics of the original trick and extensions of it, including those by Simonson & Holm [SH03], Kleber [Kle02], and Silva et al. [FHSS16, p. 147-157].

For the purposes of this paper, it is sufficient to know that the trick relies on the assistant's ability to use the four visible cards to signal three pieces of information to the magician: suit of the hidden card, a certain starting rank, and how much to add to that starting rank to obtain the rank of the hidden card. The suit and starting rank are both signaled by one card which we refer to in this paper as the *key card*. In the original trick, the key card is always placed in position 1 (farthest left). The other three visible cards are used to indicate the amount added to the rank of the key card in order to get the rank of the hidden card. A common method is to establish an ordering on the cards

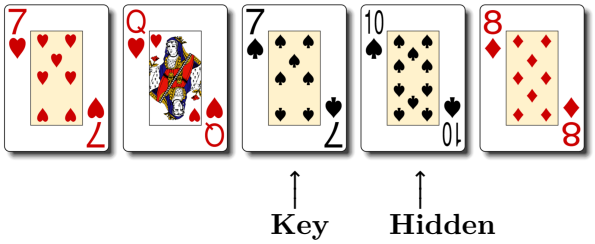
©2026 Neil J. Calkin, Eliza Gallagher, and Ben Gobler. This is an open access article licensed under the Creative Commons Attribution-NonCommercial-NoDerivs License (https://creativecommons.org/licenses/by-nc-nd/4.0/).

Recreational Mathematics Magazine **13** (23), pp. 111–128
DOI 10.2478/rmm-2026-0013

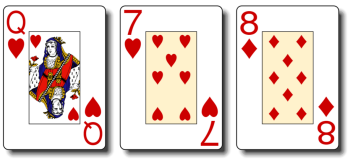
using *CHaSeD order*: clubs ♣, hearts ♡, spades ♠, diamonds ♢. The three cards put in order are assigned labels A, B, C from left to right. The six permutations of ABC are then ordered lexicographically and signal how many ranks to count up from that of the key card:

A B C → count up 1
A C B → count up 2
B A C → count up 3
B C A → count up 4
C A B → count up 5
C B A → count up 6

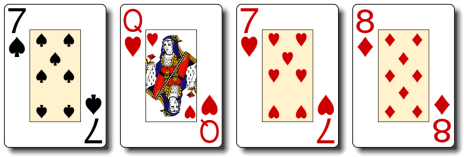
As an example, we consider the following set of five cards, which we will revisit throughout the paper. The assistant sees two cards with the same suit, 7♠ and 10♠, and decides to use them as the hidden card and key card. Since 10 is three ranks above 7, the assistant hides 10♠ and uses 7♠ as the key card.



The remaining three cards in CHaSeD order are 7♡, Q♡, and 8♢. Since the rank of the hidden card is three greater than that of the key card, the magician will need to count up 3, so the cards are arranged in BAC order:



We call these the *permuted cards*. The order of the permuted cards will be the same throughout this paper; variations of the trick will only affect the placement of the key card within the hand of four cards. To aid the gentle reader in remembering which is the key card, we note that in this example it is the sole black card among the four visible cards. For the original trick, in which the key card is placed in position 1, the four visible cards are presented to the magician in the arrangement shown here:



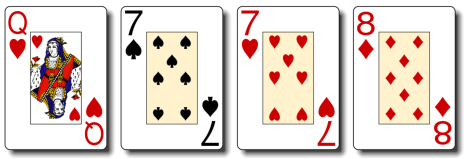
Knowing that the key card is in position 1, the magician uses the information provided by the key card and permuted cards to deduce that the hidden card is 10♠.

As it stands, the trick is impressive and can even be repeated for the same audience. But a careful observer, given the opportunity to witness the trick in its original form more than once or twice, may notice that the card in position 1 always matches the suit of the hidden card. Variety is the spice of life, and also of magic, so in this paper we play around with different conventions for relocating the key card in one of the four available positions, counted from 1 to 4, from left to right. Regardless of where the key card is positioned, the order of the permuted cards is determined simply by excluding the key card. It remains then to choose a method for selecting the position of the key card. The magician and assistant must agree ahead of time (and then remember!) which method they will use in performing the trick. Since any visible method (such as offsetting the key card or turning it face-down) would ruin the magic, these methods are mathematical in nature, all relying on arithmetic modulo 4.

The purpose of this paper is not about improving the trick, but rather about exploring some playful combinatorial ideas and analyzing the mathematics in variations of the trick. We hope that along the way we might inspire some interesting exercises for an introductory course in combinatorics.

2 Signaling with a Sum

A natural way for the assistant to relocate the key card is to begin by summing the rank values of the four visible cards, including both the key card and the permuted cards in the sum, S. The key card is then placed in position k, where $S \equiv k \pmod 4$, and we interpret the value $0 \pmod 4$ as indicating position 4. This variation is due to Colm Mulcahy [Mul03]. For our sample hand, since $7 + 12 + 7 + 8 = 34 \equiv 2 \pmod 4$, we would arrange the cards as shown, with the key card in position 2 and the permuted cards in BAC order filling the remaining positions.



Using this camouflaging technique, the position of the key card is nearly evenly distributed when considered as a random variable on all possible hands of four cards. This is not entirely obvious since the distribution of card ranks to equivalence classes modulo 4 is uneven: indeed, there are 16 cards (ranks A, 5, 9, and K) in the equivalence class of $1 \pmod 4$ while there are only 12 cards in each of the other equivalence classes.

Determining the probability of placing the key card in position k is then a calculation of which combinations of equivalence classes sum to $k \pmod 4$ and the probability of obtaining each of those combinations. A complete table of those calculations is provided in Appendix A. We note that in these calculations, as with the others discussed in this

paper, we assume that the four visible cards are selected at random from the 52-card deck. This is not entirely accurate, as it ignores the effect of the assistant choosing a nonrandom card to hide during the first stage of the trick. Nonetheless, to four digits of accuracy, expressing probabilities as percentages, we obtain:

$$\begin{aligned} p_1 &= \Pr(\text{position is } 1) = 25.01\% \\ p_2 &= \Pr(\text{position is } 2) = 24.99\% \\ p_3 &= \Pr(\text{position is } 3) = 25.01\% \\ p_4 &= \Pr(\text{position is } 4) = 24.99\% \end{aligned}$$

These probabilities are remarkably near-uniform given the simplicity of the method. However, at least one of the authors of this paper is prone to error when using this method since determining the sum may include mental conversion of face cards to their rank values of 11, 12, or 13. The next variation reduces this cognitive load, allowing for easier performance of the trick.

As for the distribution, we might also hope for a variation which gives a perfectly uniform distribution for the position of the key card. Sadly, this hope is dashed by the fact that there are $\binom{52}{4} = 270{,}725$ hands of four cards, which cannot be divided equally into four sets.

3 Signaling Using an Easier Sum

We can simplify the calculation from the previous variation by valuing all face cards at 10. Using this convention, the sum of our sample hand is $7+10+7+8 = 32 \equiv 0 \pmod 4$, so the key card would be placed in position 4.

This variation eliminates the mental conversion of face cards but also changes the probability distribution for the position of the key card. There are now 24 cards (those with ranks 2, 6, 10, J, Q, K) in the equivalence class of $2 \pmod 4$, 12 cards (those with ranks A, 5, 9) in the equivalence class of $1 \pmod 4$, and 8 cards in each of the other two equivalence classes. Calculating the probability of each key card placement uses the same process as in the previous variation but with different probabilities for the combinations (see Appendix A). Though the allotment of cards to equivalence classes is now quite imbalanced, the probability distribution for the position of the key card is only slightly less uniform than when using the actual rank values of the face cards.

$$\begin{aligned} q_1 &= \Pr(\text{position is } 1) = 24.66\% \\ q_2 &= \Pr(\text{position is } 2) = 24.74\% \\ q_3 &= \Pr(\text{position is } 3) = 25.46\% \\ q_4 &= \Pr(\text{position is } 4) = 25.14\% \end{aligned}$$

4 Signaling Using an Extreme Value (Rank)

A much simpler calculation involves no addition at all: use the value modulo 4 of the card with the lowest (or highest) rank value among the set of four visible cards. Using the lowest rank convention with our sample hand, since the lowest rank is $7 \equiv 3 \pmod 4$, we would place the key card in position 3.

Unfortunately, this method skews the probability distribution of the key card position considerably. To see this, we count the number of four-card hands in which the smallest rank is k:

$$H(k) = \binom{52 - 4(k-1)}{4} - \binom{52 - 4k}{4}$$

Then the number of four-card hands that place the key card in position i is H_i, given by

$$\begin{aligned} H_1 &= H(1) + H(5) + H(9) + H(13) \\ H_2 &= H(2) + H(6) + H(10) \\ H_3 &= H(3) + H(7) + H(11) \\ H_4 &= H(4) + H(8) + H(12) \end{aligned}$$

which yields the following probabilities for the key card position:

$$\begin{aligned} r_1 &= \Pr(\text{position is } 1) = 37.72\% \\ r_2 &= \Pr(\text{position is } 2) = 27.94\% \\ r_3 &= \Pr(\text{position is } 3) = 20.18\% \\ r_4 &= \Pr(\text{position is } 4) = 14.16\% \end{aligned}$$

The analysis for using the highest rank card is almost identical and simply permutes the above probabilities associated with each position as follows: the probabilities for positions 1 and 3 are unchanged, while the probabilities for positions 2 and 4 are interchanged. The proof of this is left as an exercise for enthusiastic readers. For either method, the heavily simplified calculation for key card position is at the cost of a distribution which is much farther from uniform than those of the previous variations.

5 Signaling Using an Extreme Value (Suit)

The simplicity of using a rank of extreme value is appealing, but the skewed distribution is unsatisfying. We can keep the one-card simplicity but improve the distribution using suits, as follows. Find the card(s) of least rank among the four cards. If this is a single card, use its suit to determine the position of the key card: ♣ = 1, ♡ = 2, ♠ = 3, ♢ = 4. If there are multiple cards of least rank, use the earliest appearing suit in CHaSeD order. In our sample hand, there are two cards with lowest rank: 7♠ and 7♡. Since hearts comes before spades in CHaSeD order, we use hearts to determine the position of the key card, placing it in position 2.

If there is a unique card of least rank, then each suit (and thus each position) is equally likely. On the other hand, if there are multiple cards with least rank, the probabilities will skew toward lower positions. But by how much? Before continuing the analysis, we note that since all cards in the deck are distinct, any collection of cards having the same rank must have distinct suits.

If there are two cards of least rank, then there are six possible pairs for their suits: ♣♡, ♣♠, ♣◇, ♡♠, ♡◇, ♠◇. Three pairs have ♣ as the least suit, two have ♡, and one has ♠. Hence, the probability of signaling position 1 is 1/2, the probability of signaling position 2 is 1/3, and the probability of signaling position 3 is 1/6 (and none of the pairs signals position 4).

If there are three cards of least rank, then there are four possible triples for their suits: ♣♡♠, ♣♡◇, ♣♠◇, ♡♠◇. In three of these the least suit is ♣, and in the other the least suit is ♡. Hence, the probability of signaling position 1 is 3/4, and the probability of signaling position 2 is 1/4 (and none of the triples signals position 3 or 4).

Finally, if there are four cards of least rank, then every suit is present, so the least suit is ♣ and the signaled position is 1.

It remains to determine the probabilities with which these four cases occur. Let x_i be the probability of having exactly i cards of lowest rank. Then

$$\begin{aligned} x_1 &= \frac{4}{\binom{52}{4}} \sum_{k=1}^{12} \binom{4k}{3} \\ x_2 &= \frac{6}{\binom{52}{4}} \sum_{k=1}^{12} \binom{4k}{2} \\ x_3 &= \frac{4}{\binom{52}{4}} \sum_{k=1}^{12} \binom{4k}{1} \\ x_4 &= \frac{1}{\binom{52}{4}} \sum_{k=0}^{12} \binom{4k}{0} \end{aligned}$$

Together with the analysis above, this implies that

$$\begin{aligned} s_1 &= \Pr(\text{position is } 1) = 28.03\% \\ s_2 &= \Pr(\text{position is } 2) = 25.93\% \\ s_3 &= \Pr(\text{position is } 3) = 23.95\% \\ s_4 &= \Pr(\text{position is } 4) = 22.09\% \end{aligned}$$

As expected, the distribution is skewed toward lower positions, but it greatly improves the distribution from the previous variation for a marginal increase in performance difficulty.

6 Summing the Suits

The final method that we will consider is the following: sum the suits of the visible cards, and use the result modulo 4 as the position of the key card. As before, we will convert suits to values by ♣ = 1, ♡ = 2, ♠ = 3, ♢ = 4. Returning once more to our sample hand, there are two ♡, one ♠, and one ♢. Since $2 + 2 + 3 + 4 = 11 \equiv 3 \pmod 4$, the key card would be placed in position 3.

To find the distribution of the key card position, we first consider the possible ways of partitioning the four visible cards into groups of suits, which we call the *shape* of the hand. The numbers of hands with each shape are counted as follows:

Shape $\mathbf{A_4}$: 4 of the same suit: $\binom{4}{1}\binom{13}{4} = 2{,}860$

Shape $\mathbf{A_{31}}$: 3 of one suit, 1 of another: $\binom{4}{1}\binom{13}{3}\binom{3}{1}\binom{13}{1} = 44{,}616$

Shape $\mathbf{A_{22}}$: 2 of one suit, 2 of another: $\binom{4}{2}\binom{13}{2}\binom{13}{2} = 36{,}504$

Shape $\mathbf{A_{211}}$: 2 of one suit, 1 each of two other suits: $\binom{4}{1}\binom{13}{2}\binom{3}{2}\binom{13}{1}^2 = 158{,}184$

Shape $\mathbf{A_{1111}}$: 1 of each suit: $\binom{13}{1}^4 = 28{,}561$

Gentle readers (of which we hope there will be more than one) can be reassured by the fact that these counts do indeed sum to $\binom{52}{4}$. Next, we can determine the distribution of key card positions among hands of each shape. We find the following:

Shape $\mathbf{A_4}$: the position is always 4

Shape $\mathbf{A_{31}}$: the distribution is uniform on positions 1, 2, 3

Shape $\mathbf{A_{22}}$: the position is 2 with probability 2/3 and 4 with probability 1/3

Shape $\mathbf{A_{211}}$: the distribution is uniform on 1, 3, 4

Shape $\mathbf{A_{1111}}$: the position is always 2

To demonstrate one of these calculations, we consider the hands of Shape A_{22}, which have the following sums and hence give the distribution listed above:

Pairs of Suits	Sum Modulo 4
♣♣ ♡♡	2
♣♣ ♠♠	0
♣♣ ♢♢	2
♡♡ ♠♠	2
♡♡ ♢♢	0
♠♠ ♢♢	2

Using the distributions for each hand shape, we can compute the total number of hands that signal each key card position:

$$\begin{aligned}
\text{Position 1:} \quad & \tfrac{1}{3}A_{31} + \tfrac{1}{3}A_{211} = 67{,}600 \\
\text{Position 2:} \quad & \tfrac{1}{3}A_{31} + \tfrac{2}{3}A_{22} + A_{1111} = 67{,}769 \\
\text{Position 3:} \quad & \tfrac{1}{3}A_{31} + \tfrac{1}{3}A_{211} = 67{,}600 \\
\text{Position 4:} \quad & A_4 + \tfrac{1}{3}A_{22} + \tfrac{1}{3}A_{211} = 67{,}756
\end{aligned}$$

Dividing each total by $\binom{52}{4}$, we get the corresponding probabilities:

$$\begin{aligned}
u_1 &= \Pr(\text{position is } 1) = 24.97\% \\
u_2 &= \Pr(\text{position is } 2) = 25.03\% \\
u_3 &= \Pr(\text{position is } 3) = 24.97\% \\
u_4 &= \Pr(\text{position is } 4) = 25.03\%
\end{aligned}$$

So, while the shapes of hands are very much *not* uniformly distributed, and our intuition may lead us to believe that the distribution of key card positions should therefore be far from uniform, it turns out that the convention of summing the suits does indeed get very close to a uniform distribution.

7 Independence Makes Everything Better

It is unlikely that an audience member will catch on to any of the variations that we have mentioned so far, even when the trick is performed multiple times. But nothing stops us from inventing ones that are far more contrived. For instance, we can combine any set of variations as follows: calculate the position of the key card according to each variation, and sum the results modulo 4. Using our sample hand, the previous two methods placed the key card in positions 2 and 3, respectively. Since $2 + 3 \equiv 1 \pmod 4$, the amalgam method would place the key card in position 1. However, the positions signaled by different methods are *dependent* random variables. For instance, a hand with highest rank 2 cannot have shape A_{31}.

In fact, dependence has been present in every method involving sums: since the cards in a hand are distinct, each appearing card restricts the possibilities for the other cards. The calculations in this paper have been modest, but we draw the line at the amalgam method mentioned above, for which dependence makes the casework considerably more challenging.

In this section we'll see how independence makes for simpler calculations and can improve the uniformity of the distribution for key card positions. Independence is achieved by using information from multiple performances of the trick. For example, if the signaling convention were to sum the least suits of the last two performances and use that value modulo 4 as the key card position, then we are summing independent random variables. Pleasantly, this convention gives a more uniform distribution than that of the least suit method, which was used for the previous performances individually. In general, we have the following result:

Proposition 1. *Suppose that X and Y are two independent non-uniform random variables on $\{0,1,2,3\}$, and $Z = X + Y \pmod 4$. Suppose that $Pr(X = i) = r_i$, $Pr(Y = i) = s_i$, $Pr(Z = i) = t_i$, and that for each i, $r_i > 0$ and $s_i > 0$. Then the distribution of Z is closer to uniform than each of X, Y, in the following sense:*

$$\min_i t_i > \min_i r_i \qquad \text{and} \qquad \max_i t_i < \max_i r_i$$

and similarly,

$$\min_i t_i > \min_i s_i \qquad \text{and} \qquad \max_i t_i < \max_i s_i$$

Proof.

$$t_k = \sum_{i+j \equiv k \pmod 4} r_i s_j > \min_i r_i \sum_j s_j = \min_i r_i$$

The assertions for $\max_i r_i$, $\min_i s_i$, and $\max_i s_i$ are handled similarly. □

This surprisingly brief argument shows that the sum of two such independent non-uniform distributions modulo 4 will result in a distribution that is closer to uniform than each of the original distributions. To see this in action, we visit an example.

Example 1. *Let X and Y have the following distributions on $\{0,1,2,3\}$:*

$$\begin{aligned} r_0 &= \Pr(X = 0) = 10\% & s_0 &= \Pr(Y = 0) = 20\% \\ r_1 &= \Pr(X = 1) = 20\% & s_1 &= \Pr(Y = 1) = 20\% \\ r_2 &= \Pr(X = 2) = 30\% & s_2 &= \Pr(Y = 2) = 10\% \\ r_3 &= \Pr(X = 3) = 40\% & s_3 &= \Pr(Y = 3) = 50\% \end{aligned}$$

Since X and Y are independent, the distribution of $X + Y$ is calculated as follows:

$$\begin{aligned}
\Pr(X+Y=0) &= r_0 s_0 = 2\% \\
\Pr(X+Y=1) &= r_0 s_1 + r_1 s_0 = 6\% \\
\Pr(X+Y=2) &= r_0 s_2 + r_1 s_1 + r_2 s_0 = 11\% \\
\Pr(X+Y=3) &= r_0 s_3 + r_1 s_2 + r_2 s_1 + r_3 s_0 = 21\% \\
\Pr(X+Y=4) &= r_1 s_3 + r_2 s_2 + r_3 s_1 = 21\% \\
\Pr(X+Y=5) &= r_2 s_3 + r_3 s_2 = 19\% \\
\Pr(X+Y=6) &= r_3 s_3 = 20\%
\end{aligned}$$

The distribution of $Z = X + Y \pmod 4$ is then given by

$$\begin{aligned}
t_0 &= \Pr(Z=0) = \Pr(X+Y=0 \text{ or } 4) = 23\% \\
t_1 &= \Pr(Z=1) = \Pr(X+Y=1 \text{ or } 5) = 25\% \\
t_2 &= \Pr(Z=2) = \Pr(X+Y=2 \text{ or } 6) = 31\% \\
t_3 &= \Pr(Z=3) = \Pr(X+Y=3) = 21\%
\end{aligned}$$

We see that each t_i is strictly between 10% and 40%, as expected.

What about summing the results of more than two performances? If $X_1, X_2, \ldots, X_n$ are independent identically distributed random variables with the same distribution as X, then the distribution of their sum modulo 4 converges exponentially quickly to the uniform distribution as $n \to \infty$. Ambitious readers are welcome to reference Appendix B for details on this result.

8 Conclusion

We've analyzed several schemes for signaling the position of a key card in a hand of four cards, with a focus on the underlying mathematics. In each method, there is a tradeoff between ease of calculation and uniformity of the distribution for the key card position. Magical friends are encouraged to explore and determine which balance is most pleasing to them.

References

[Dia03] Persi Diaconis. Random walks on groups: characters and geometry. In *Groups St. Andrews 2001 in Oxford. Vol. I*, volume 304 of *London Math. Soc. Lecture Note Ser.*, pages 120–142. Cambridge Univ. Press, Cambridge, 2003. doi:10.1017/CBO9780511542770.017.

[FHSS16] Pedro Jorge Freitas, Tiago Hirth, Alexandre Silva, and Jorge Nuno Silva. *Matemagia: Matemagia com cartas de jogar (Portuguese edition)*. Associação Ludus, 2016.

[Kle02] Michael Kleber. The best card trick. *The Mathematical Intelligencer*, 24(1):9–11, 2002. doi:10.1007/BF03025305.

[Lee51] W. Wallace Lee. *Math Miracles*. Calgary: Micky Hades International, 1951.

[Mul03] Colm Mulcahy. Fitch Cheney's five card trick. *Math Horizons*, 10:11–13, 2003.

[Mul13] Colm Mulcahy. *Mathematical Card Magic: Fifty-Two New Effects*. CRC Press, 2013.

[Rio68] John Riordan. *Combinatorial identities*, volume 217. Wiley New York, 1968.

[SH03] Shai Simonson and Tara S. Holm. Using a card trick to teach discrete mathematics. *Primus : Problems, Resources, and Issues in Mathematics Undergraduate Studies*, 8(3):248–269, 09 2003.

[Sim57] Thomas Simpson. Ciii. the invention of a general method for determining the sum of every 2d, 3d, 4th, or 5th, &c. term of a series, taken in order; the sum of the whole series being known. *Philosophical Transactions of the Royal Society of London*, 50:757–769, 1757.

A Calculations for Key Card Position by Summing Card Values under Two Conventions

The following tables give the counts of hands that signal each position using the sum of all four visible cards modulo 4.

Sum of All Four Cards	**Composition**				**Count**	
	[0]	[1]	[2]	[3]	rank value	face cards as 10
1(mod 4)	3	1	0	0	$16 \cdot \binom{12}{3}$	$12 \cdot \binom{8}{3}$
	2	0	1	1	$12 \cdot 12 \cdot \binom{12}{2}$	$8 \cdot 24 \cdot \binom{8}{2}$
	1	2	0	1	$12 \cdot 12 \cdot \binom{16}{2}$	$8 \cdot 8 \cdot \binom{12}{2}$
	1	1	2	0	$12 \cdot 16 \cdot \binom{12}{2}$	$8 \cdot 12 \cdot \binom{24}{2}$
	1	0	0	3	$12 \cdot \binom{12}{3}$	$8 \cdot \binom{8}{3}$
	0	3	1	0	$12 \cdot \binom{16}{3}$	$24 \cdot \binom{12}{3}$
	0	1	1	2	$16 \cdot 12 \cdot \binom{12}{2}$	$12 \cdot 24 \cdot \binom{8}{2}$
	0	0	3	1	$12 \cdot \binom{12}{3}$	$8 \cdot \binom{24}{3}$
	Total Hands				67,648	66,752
	Probability				0.249877182	0.24656755

Table 1: Probability of placing the key card in position 1 by calculating the sum of the four visible cards modulo 4, either using the rank value or counting face cards as 10.

Sum of All Four Cards	**Composition**				**Count**	
	[0]	[1]	[2]	[3]	rank value	face cards as 10
2(mod 4)	3	0	1	0	$12 \cdot \binom{12}{3}$	$24 \cdot \binom{8}{3}$
	2	2	0	0	$\binom{12}{2} \cdot \binom{16}{2}$	$\binom{8}{2} \cdot \binom{12}{2}$
	2	0	0	2	$\binom{12}{2}\binom{12}{2}$	$\binom{8}{2}\binom{8}{2}$
	1	1	1	1	$12^3 \cdot 16$	$8^2 \cdot 12 \cdot 24$
	1	0	3	0	$12 \cdot \binom{12}{3}$	$8 \cdot \binom{24}{3}$
	0	3	0	1	$12 \cdot \binom{16}{3}$	$8 \cdot \binom{12}{3}$
	0	2	2	0	$\binom{16}{2}\binom{12}{2}$	$\binom{12}{2}\binom{24}{2}$
	0	1	0	3	$16 \cdot \binom{12}{3}$	$12 \cdot \binom{8}{3}$
	0	0	2	2	$\binom{12}{2}\binom{12}{2}$	$\binom{24}{2}\binom{8}{2}$
	Total Hands				67,720	66,976
	Probability				0.250143134	0.247394958

Table 2: Probability of placing the key card in position 2 by calculating the sum of the four visible cards modulo 4, either using the rank value or counting face cards as 10.

Sum of All Four Cards	**Composition**				**Count**	
	[0]	[1]	[2]	[3]	rank value	face cards as 10
3(mod 4)	3	0	0	1	$12 \cdot \binom{12}{3}$	$8 \cdot \binom{8}{3}$
	2	1	1	0	$16 \cdot 12 \cdot \binom{12}{2}$	$12 \cdot 24 \cdot \binom{8}{2}$
	1	3	0	0	$12 \cdot \binom{16}{3}$	$8 \cdot \binom{12}{3}$
	1	1	0	2	$12 \cdot 16 \cdot \binom{12}{2}$	$8 \cdot 12 \cdot \binom{8}{2}$
	1	0	2	1	$12 \cdot 12 \cdot \binom{16}{2}$	$8 \cdot 8 \cdot \binom{24}{2}$
	0	2	1	1	$12 \cdot 12 \cdot \binom{16}{2}$	$24 \cdot 8 \cdot \binom{12}{2}$
	0	1	3	0	$16 \cdot \binom{12}{3}$	$12 \cdot \binom{24}{3}$
	0	0	1	3	$12 \cdot \binom{12}{3}$	$24 \cdot \binom{8}{3}$
	Total Hands				67,648	68,928
	Probability				0.249877182	0.254605227

Table 3: Probability of placing the key card in position 3 by calculating the sum of the four visible cards modulo 4, either using the rank value or counting face cards as 10.

Sum of All Four Cards	**Composition**				**Count**	
	[0]	[1]	[2]	[3]	rank value	face cards as 10
0(mod 4)	4	0	0	0	$\binom{12}{4}$	$\binom{8}{4}$
	2	1	0	1	$16 \cdot 12 \cdot \binom{12}{2}$	$12 \cdot 8 \cdot \binom{8}{2}$
	2	0	2	0	$\binom{12}{2}\binom{12}{2}$	$\binom{8}{2}\binom{24}{2}$
	1	2	1	0	$12 \cdot 12 \cdot \binom{16}{2}$	$8 \cdot 24 \cdot \binom{12}{2}$
	1	0	1	2	$12 \cdot 12 \cdot \binom{12}{2}$	$8 \cdot 24 \cdot \binom{8}{2}$
	0	4	0	0	$\binom{16}{4}$	$\binom{12}{4}$
	0	2	0	2	$\binom{16}{2}\binom{12}{2}$	$\binom{12}{2}\binom{8}{2}$
	0	1	2	1	$16 \cdot 12 \cdot \binom{12}{2}$	$12 \cdot 8 \cdot \binom{24}{2}$
	0	0	4	0	$\binom{12}{4}$	$\binom{24}{4}$
	0	0	0	4	$\binom{12}{4}$	$\binom{8}{4}$
	Total Hands				67,709	68,069
	Probability				0.250102503	0.251432265

Table 4: Probability of placing the key card in position 4 by calculating the sum of the four visible cards modulo 4, either using the rank value or counting face cards as 10.

B Summing Identical Independent Random Variables Modulo 4

We will consider what happens if we take many independent *identical* random variables in $\mathbb{Z}_4 = \{0, 1, 2, 3\}$ and sum them modulo 4. Suppose that we have a random variable X taking values in $\mathbb{Z}_4$ with $\Pr(X = i) = p_i$. To avoid degeneracy, we will assume that each $p_i > 0$. Write $f(x) = p_0 + p_1 x + p_2 x^2 + p_3 x^3$, called the *probability generating function* of X. Let $X_1, X_2, \ldots, X_n$ be independent identically distributed random variables with the same distribution as X, and let

$$Z = \sum_{i=1}^{n} X_i$$

be the sum (in $\mathbb{Z}$) of the X_i's: we wish to determine the distribution of Z (mod 4).

The distribution of Z is given by the coefficients of the polynomial $g(x) = f(x)^n$. To obtain the distribution of Z (mod 4), we must separate out $g(x)$ as $g(x) = g_0(x) + g_1(x) + g_2(x) + g_3(x)$, where

$$\begin{aligned} g_0(x) &= \sum_{k\geq 0} p_{4k} x^{4k} \\ g_1(x) &= \sum_{k\geq 0} p_{4k+1} x^{4k+1} \\ g_2(x) &= \sum_{k\geq 0} p_{4k+2} x^{4k+2} \\ g_3(x) &= \sum_{k\geq 0} p_{4k+3} x^{4k+3} \end{aligned}$$

The distribution of Z (mod 4) is then given by the values of $g_0(1)$, $g_1(1)$, $g_2(1)$, $g_3(1)$. This type of filtering for every fourth term in a series is an example of what is called a *series multisection*. The technique for finding closed-form representations of series multisections is due to Thomas Simpson in the 18th century and involves primitive roots of unity [Sim57, Rio68, p. 131-141]. Many readers will be familiar with the form of this method for series bisection into even and odd terms. The general form is not too much different. Here are the standard results applied to each $g_i(x)$:

$$\begin{aligned} g_0(x) &= \frac{1}{4}\left(g(x) + g(ix) + g(-x) + g(-ix)\right) \\ g_1(x) &= \frac{1}{4}\left(g(x) - ig(ix) - g(-x) + ig(-ix)\right) \\ g_2(x) &= \frac{1}{4}\left(g(x) - g(ix) + g(-x) - g(-ix)\right) \\ g_3(x) &= \frac{1}{4}\left(g(x) + ig(ix) - g(-x) - ig(-ix)\right) \end{aligned}$$

where i is the imaginary unit (and remarkably, the imaginary components will cancel out, leaving real-valued coefficients). What happens when n gets large?

We remind the gentle reader that since the p_i's denote a probability distribution, they must sum to 1, so it follows that $f(1) = 1$. And $f(-1) = p_0 + p_2 - p_1 - p_3 < 1$ since all of the p_i's are positive. Also, $f(i) = (p_0 - p_2) + (p_1 - p_3)i$ and $f(-i) = (p_0 - p_2) - (p_1 - p_3)i$, so $|f(i)| = |f(-i)| < 1$. Hence, each of $f(-1), f(i), f(-i)$ have magnitude less than 1. Since $g(x) = f(x)^n$, it follows that $g(-1), g(i), g(-i)$ all converge exponentially quickly to zero as $n \to \infty$. So, $g_0(1), g_1(1), g_2(1), g_3(1)$ all converge exponentially quickly to $\frac{1}{4}g(1) = \frac{1}{4}$. That is, the distribution of $Z \pmod 4$ converges to the uniform distribution.

It can be seen that a similar result holds for products over any finite group. Let $G = \{g_1, g_2, \dots g_N\}$ be a finite group with identity $e = g_1$. Associate with each group element g_i a probability p_i. If we select elements $h_1, h_2, \dots, h_n$ independently according to this distribution and compute the random variable $X_n = h_1 \circ h_2 \circ \cdots \circ h_n$, how does $\Pr(X_n = g_i)$ behave as $n \to \infty$? These sorts of random walks on G have been studied extensively by Diaconis [Dia03] and others. Just as in the case where $G = (\mathbb{Z}_4, +)$, if we assume that each $p_i > 0$ to avoid degeneracy, then the distribution of X_n converges to the uniform distribution.

To see why, let $v^{(n)}$ be a vector of probabilities in $\mathbb{R}^N$ where the ith entry of $v_i^{(n)}$ is the probability that $X_n = g_i$. Define permutation matrices $M_1, \dots M_N$ corresponding to the permutations on the elements of G induced by multiplication on the right by $g_1, \dots g_N$, respectively. Let

$$M = p_1 M_1 + p_2 M_2 + \cdots + p_N M_N$$

Then the rows (and columns) of M are permutations of the p_i's. Since each $p_i > 0$, we have that M is a positive matrix. This is the sole condition of the Perron–Frobenius theorem, which states that such a matrix M has a unique eigenvalue of maximum absolute value, and this eigenvalue is positive and corresponds with the only eigenvector with all positive entries. Notice that the vector $v = (1/N, 1/N, \dots, 1/N)$ is an eigenvector of M with eigenvalue 1. Hence, the largest eigenvalue of M is 1, and all other eigenvalues of M have magnitude strictly less than 1.

Finally, since $v^{(n)} = M^n v^{(0)}$, we conclude that $v^{(n)}$ converges into the eigenspace spanned by the eigenvector v as $n \to \infty$. Moreover, since $v^{(n)}$ is a vector of probabilities summing to 1, it must then converge to v. The rate of convergence is exponential, determined by the size of the second largest eigenvalue of M. Thus, in the general setting we have exponential convergence to the uniform distribution.

We can relax the assumption that each $p_i > 0$ as long as the following two conditions are met:

(i) the probabilities are positive on a set of elements which generate G, and

(ii) the probabilities do not induce periodic behavior on the powers of M (for example, in $\mathbb{Z}_4$, having $\Pr(h = 1) = 1$ and all other probabilities 0).

Given (i), a simple way to satisfy (ii) is to have $\Pr(h = e) > 0$. These two conditions ensure that some power of M is positive, and hence the Perron–Frobenius theorem applies and gives the results above.