DETERMINISTIC AND PROBABILISTIC APPROACHES TO CARD SHUFFLING

CASSIDY AMASON

ABSTRACT. In this paper we consider the riffle shuffle and analyze the outcomes of deterministic and probabilistic shuffles. Famous sleight-of-hand experts Persi Diaconis and Martin Gardner have contributed widely to this field of study. We first show how to move an individual card to any given position within the deck. This being a matter of skill, we then move to the more common "clumsy" shuffling. Throughout this paper we analyze the properties associated with the group of all permutations for a deck of 52 cards and ultimately determine how much shuffling is required to randomize a deck.

1. Introduction

The ability to manipulate the probabilities in a deck of cards through a simple choice in shuffle is a concept known to few. The goal of shuffling is to create a uniform distribution among the deck of cards. Without the deck's being randomized, an unfair advantage is brought out. Persi Diaconis and Martin Gardner are famous sleight-of-hand magicians who have devoted their time and energy to understanding the rules of probability to manipulate a deck of cards so illusions may take place. These illusions are then interpreted as magic.

We will first take a look at types of faro shuffles, or perfect riffle shuffles. Using these types of shuffles, we can determine how many shuffles are required to restore a deck of even or odd number to its natural order. We may also calculate the number of shuffles required to move individual cards to any given position. If the shuffles are not performed perfectly, the desired outcome becomes unachievable. Thus, these types of shuffles require an ample amount of skill. We then consider the more common approach to shuffling - clumsy shuffling. Using this approach, we can then analyze rising sequences, as well as convolving densities on S_{52} , to ultimately discovering how many shuffles

Date: November 30, 2016.

may randomize a deck of 52 cards.

Famous sleight-of-hand expert Persi Diaconis was born in 1945, and at the age of 5, he developed a deep interest in magic. He began to teach himself simple tricks that continued to increase in sophistication over time. Diaconis attended George Washington High School in NYC and found himself at home as a member of the magic club. Often, Diaconis would cut school and hang around magic stores where he learned about different types of shuffles. As Diaconis' education level grew, he was able to use math in many of his card tricks. Before Diaconis was able to gradutate at the age of 15, Dai Vernon (a famous canadian born sleight-of-hand expert) invited 14 year old Diaconis to join him on one of his American tours performing magic shows. Diaconis eagerly left without a word to his parents. Regardless of not being in high school, Diaconis' teachers decided to give him grades for exams he had not taken - and he ended up graduating high school. Diaconis started out his college career at City University in NY, and paid his way through college collecting money he made from playing poker games on ships. Diaconis eventually met Martin Gardner because of their shared interest in magic, and Gardner made arrangements for Diaconis to attend Harvard University, where he majored in statistics. He started inventing magic tricks, giving lessons, and collecting old books on magic and at 71 he is still devoting much of his energy to magic. Diaconis is currently a Professor of Statistics and Math at Stanford University. [9]

Born in 1914, Martin Gardner came from a wealthy family who owned a small oil company. He attended the University of Chicago where he majored in philosophy. Over the course of his life, Gardner published many works covering an astonishing number of fields. His first published work was a magic trick for the Sphinx at the age of 15. Initially, he wrote little pieces for magazines that didn't pay. After serving in World War II for four years, he wrote for Esquire Magazine, was laid off, and then moved to writing for Humpty Dumpty (a children's magazine). Gardner incorporated a number of paper cuttings and foldings into the magazine. Eight years later, he went to write for Scientific American, writing a column titled Mathematical Games. He has written over 60 hardbacks and a number of mathematical papers - a few of which we reference in this paper. In 2010, at the age of 95, Martin Gardner passed away. [8]

2. Deterministic Magic

In this section, we introduce a few perfect shuffles in which we are able to watch individual cards move throughout the deck with certainty.

Definition 2.1. A riffle shuffle is done by splitting a deck of N cards into 2 piles of k and N - k and interlacing the piles together.

The shuffles that are defined below are examples of perfect riffle shuffles.

Definition 2.2. An faro shuffle is a perfect riffle shuffle, in which the deck is split evenly and the two halves are interwoven perfectly.

Some types of faro shuffles include in-shuffles and out-shuffles.

Definition 2.3. An out-shuffle is a faro shuffle in which the original top card remains on top after the shuffle has ended.

Definition 2.4. An *in-shuffle* is a faro shuffle in which the original top card takes the position of the second card from the top after the shuffle has ended.

Using the piece-wise functions below, we are able to determine where any card k in a deck of N cards will move after a single in-shuffle

$$F_I(k) = \begin{cases} 2k & k \le N/2 \\ 2(k-N/2)-1 & k > N/2 \end{cases}$$

and after a single out-shuffle

$$F_O(k) = \begin{cases} 2k-1 & k \le N/2\\ 2(k-N/2) & k > N/2. \end{cases}$$

A faro shuffle works well with an even number of cards because we can split the deck evenly in half, but what about odd numbered decks? For an odd deck of cards, if the deck is cut above the center card it is an in-shuffle. This places the top card on the smaller half, resulting with its being second from the top. If the deck is cut below the center card, it is an out-shuffle. This places the top card on the larger half, resulting in the top card remaining in first position.

Lemma 2.5. Let N be the number of cards in a deck. Then repeated faro shuffles of the same type will return the deck to its original order.

If N is odd, then x number of in or out-shuffles will return the deck to its original order, where x is in the formula: $2^x = 1(modN)$. If N is even, then x number of out-shuffles will return the deck to its original order, where x is in the formula: $2^x = 1(mod(N-1))$ or x number of in-shuffles will return the deck to its original order, where x is in the formula: $2^x = 1(mod(N+1))$. Note that for each even N, the number of out-shuffles to restore order will be the same number of shuffles as N-1.

Example: Take N=51 and N=52. For N=51, we note 51 is odd. Observe the following:

$$2^6 = 64 = 13 \pmod{51}$$

 $2^7 = 128 = 26 \pmod{51}$
 $2^8 = 256 = 1 \pmod{51}$.

Thus, 8 in-shuffles or 8 out-shuffles return the deck to its original order. Next we take N=52, and we note 52 is even. Above we observed that for each even deck, the number of out-shuffles to restore order will be the same number of shuffles as the odd numbered deck before it (N-1). Thus, 8 out-shuffles restore an even deck of N=52 to original order. To find the number of in-shuffles, observe the following:

$$2^7 = 128 = 22 \pmod{(52+1)}$$

$$2^8 = 256 = 44 \pmod{(52+1)}$$

$$\vdots$$

$$\vdots$$

$$2^{52} = 4.50 * 10^{15} = 1 \pmod{(52+1)}$$

Thus, for N = 52, 8 out-shuffles or 52 in-shuffles will return the deck to original order. In [4], Gardner gives a chart that depicts the number of in and out shuffles required to restore order for a deck of 2 cards to a deck of 52 cards. The following is a small portion of Gardner's chart:

Number of Cards in Deck	Out-Shuffles	In-Shuffles
7	3	3
8	3	6
9	6	6
50	21	8
51	8	8
52	8	52

Lemma 2.6. There is a method of shuffling in which we can bring the top card of a deck to any given position within the deck.

Let a deck of N cards be labeled 1, 2, 3, ..., N. To bring the top card to position p we take p-1=y and write y as a base 2 number to obtain a representation consisting of 1's and 0's. Reading from left to right, we can then associate the 1's to in-shuffles and the 0's to out-shuffles.

Example: Say we have a deck of 8 cards and we want to move the top card to position 5. Then, N=8 and p=5. Note that 5-1=4. In base 2, we have 4=100. Thus, after assigning in-shuffles to the 1's and out-shuffles to the 0's, we have that one in-shuffle followed by two out-shuffles will move the top card to position 5.

Observe the following:

Suppose our deck starts in natural order: [A 2 3 4 5 6 7 8].

After one in-shuffle: [5 A 6 2 7 3 8 4] After an out-shuffle: [5 7 A 3 6 8 2 4] After the last out-shuffle: [5 6 7 8 A 2 3 4]

In [6], Paul Swinford, a card shuffling expert, discovered that the pattern of shuffles that brings the top card to a given position p also brings the card in position p to the top, as seen in our previous example. However, if we take a card in position p and move it to the top, it is not always true that the top card will move to position p.

Lemma 2.7. There is a method of shuffling in which we can bring a card in any position within the deck to the top.

Let a deck of 2n cards be labeled 0, 1, 2, 3, ..., 2n-1, with the top card being in position 0. Let $r \in \mathbb{Z}$ satisfying $2^{r-1} < 2n \le 2^r$. We can assume that 0 since if <math>p=0 we don't have to do anything and if p=2n-1, then r consecutive in-shuffles does the job. Next, let t be the largest integer not exceeding $\frac{2^r(p+1)}{2n}$ and write $t=t_1,t_2,...t_r$ it its

base 2 expansion. Let $s_1, s_2, ...s_r$ be the last r digits of the base expansion of 2nt. Finally, form the binary sequence $u_1, u_2, ...u_r$ by defining $u_i = s_i + t_i$ where addition is done modulo 2. Then the desired shuffle sequence can be read from left to right associating the 1's to in-shuffles and the 0's to out-shuffles.

Example: Let 2n = 12, and we want to bring card in position 8 to the top.

First, r=4 since $2^3<12\leq 2^4$. Then, $\frac{2^4(8+1)}{12}=12$. Thus t=11. In base 2, we have 11=1011. Then, 2nt=12*11=132=10000100, and the last r=4 digits $=0100=S_i$. Finally, $u_i=s_i+t_i=0100+1011=1111$. Thus, the resulting shuffle sequence to bring card in position 8 to the top is four consecutive in-shuffles.

Observe the following:

Suppose our deck starts in natural order: [A 2 3 4 5 6 7 8 9 10 J Q]. Note that card A is in position 0 and the card in position 8 is card 9. After the first in-shuffle we have: [7 A 8 2 9 3 10 4 J 5 Q 6] After the second in-shuffle: [10 7 4 A J 8 5 2 Q 9 6 3] After the third in-shuffle: [5 10 2 7 Q 4 9 A 6 J 3 8] And finally, after the fourth in-shuffle: [9 5 A 10 6 2 J 7 3 Q 8 4]. Note that while the card in position 8 came to the top, our original top card did not move to position 8.

Moving individual cards within a deck with certainty takes skill - it is estimated that fewer than 100 people in the world can do these types of shuffles with a 52 card deck perfectly. Most people are clumsy shufflers, so now we consider a probabilistic approach to shuffling and magic.

3. Background for Shuffling

Definition 3.1. A shuffle or permutation of cards is a rearrangement of cards (one-to-one and onto) $\pi : \{1, ..., N\} \rightarrow \{1, ..., N\}$.

We let S_N denote all permutations of the set $\{1, 2, 3, ..., N\}$. So for S_{52} (we are denoting the set of all permutations for a deck of 52 cards), the possible number of rearrangements is 52!. Thus, any arrangement of the deck is an element of S_{52} , and shuffling can be compared to S_{52} acting on itself. The specifics of the shuffle are determined by the

probability distribution of shuffles, π .

Ideally, we want every arrangement to have probability $\frac{1}{N!}$ for N cards, such that adding each probability distribution will equal 1. This means the deck has a uniform distribution - in other words, the deck is completely randomized.

If we were to look at a deck before and after a shuffle, how many cards would still be in the same position? Knowing these probabilities are useful for card dealers and magicians when performing certain tricks. First, we choose a completely random shuffle on S_N . A completely random shuffle will be a permutation chosen from S_N with each permutation having probability $\frac{1}{N!}$ of being chosen.

Take for example N=2. There are 2!=2*1=2 permutations that can result from shuffling 2 cards. Take cards A and 2. The permutations are as follows: $\{[A2],[2A]\}$. Let X= the number of k cards in the same position. Then:

k	0	1	2
P(X=k)	1/2	0	1/2

Thus the expected number of cards that stay in the same position are

$$E(X) = 0(1/2) + 1(0) + 2(1/2) = 1$$

and the probability that at least 1 card stays in the same position is

$$P(X > 1) = 1 - (1/2) = .5.$$

Next we take N=3. There are 3!=3*2*1=6 permutations that can result from shuffling 3 cards. Take cards A,2,3. The permutations are as follows: {[A23],[A32],[2A3],[23A],[32A],[3A2]}. If we again let X = the number of k cards in the same position, then

k	0	1	2	3
P(X=k)	2/6	3/6	0	1/6

The expected number of cards that stay in the same position is

$$E(X) = 0(2/6) + 1(3/6) + 3(1/6) = 1$$

and

$$P(X > 1) = 1 - (2/6) = .667.$$

If we wanted to know how many cards stay in the same position for a standard deck of cards, it would be impossible to list all 52! permutations and count which cards stay fixed for each permutation. Thus, we may instead observe the derangements in S_{52} .

Definition 3.2. A **derangement** is a permutation of a set that leaves no element in its original state.

While N! denotes the number of possible rearrangements in S_N , !N (the subfactorial of N) denotes the possible derangements in S_N . The subfactorial of N is given by the formula:

$$!N = N!(1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \dots + \frac{(-1)^N}{N!})$$

and for S_{52} ,

$$52!(1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \dots + \frac{(-1)^{52}}{52!})$$

or

$$52(\sum_{N=0}^{52} \frac{-1^N}{N!}).$$

Note that from the Maclaurin series we are given

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

and

$$e^{-1} = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!}.$$

Thus, the probability that one or more elements stay fixed after a single shuffle is $1 - \frac{52!/e}{52!}$, which is approximately 63%.

4. Convolutions

Repeatedly shuffling a deck of cards corresponds to composing shuffles within the deck. Let a be a particular arrangement of N cards in the set all arrangements of N cards (S_N) . If we pick a shuffle (π) and apply it to a, we acquire a new arrangement (a') that is still an element S_N , thus S_N acts on itself. Say we start with an arrangement a=[123456] and the shuffle we pick is $\pi=(1432)$. Observe the following chart:

$$\begin{array}{c|cc} a & [1 \ 2 \ 3 \ 4 \ 5 \ 6] \\ \hline \pi(a) = a' & [2 \ 3 \ 4 \ 1 \ 5 \ 6] \\ \pi(a') = a'' & [3 \ 4 \ 1 \ 2 \ 5 \ 6] \end{array}$$

What is the probability that we chose $\pi = (1432)$ as our shuffle? Let Q denote the step distribution, or the probability of choosing $\pi \epsilon S_N$ as the shuffle.

We will use the top-in shuffle to illustrate our next example. A top in shuffle can be done by simply taking the top card of the deck and inserting it anywhere into the deck. Thus, observe the only following rearrangements for N=6:

$$\begin{aligned} \pi_1 &= (1) = [123456] \\ \pi_2 &= (12) = [213456] \\ \pi_3 &= (132) = [231456] \\ \pi_4 &= (1432) = [234156] \\ \pi_5 &= (15432) = [234516] \\ \pi_6 &= (165432) = [234561] \end{aligned}$$

Hence, the probability of choosing one of the above shuffles (Q) is 1/6.

Definition 4.1. A random walk on the symmetric group (S_N) is a sequence of random shuffles each chosen from the same probability distribution.

In [7], Mann tells us to suppose we are given a method of shuffling. We start at the identity permutation, S_i and take a step in the random walk (choose a π_1) and then a second step in the random walk, π_2 . Performing π_1 and then π_2 to a deck of cards can be written as $\pi_2 \circ \pi_1$. What is the probability for any given permutation now? In other words, what is the density for $\pi_2 \circ \pi_1$? We may call this density $Q^{(2)}$. Note that the probability for π_1 and then π_2 being chosen is $Q(\pi_1) * Q(\pi_2)$, since the choices are independent of each other. So for any permutation, π , $Q^{(2)}(\pi)$ is given by the sum of $Q(\pi_1) * Q(\pi_2)$ such

that $\pi = \pi_2 \circ \pi_1$.

This way of combining Q with itself is called a convolution and is written Q * Q:

$$Q^{(2)}(\pi) = Q * Q(\pi) = \sum_{\pi_2 \circ \pi_1 = \pi} (Q(\pi_1)Q(\pi_2)) = \sum_{\pi_1} (Q(\pi_1)Q(\pi \circ \pi_1^{-1}))$$

Lemma 4.2. The associative property holds for the convolutions of 3 probability distributions, p. In other words, $(p_1 * (p_2 * p_3))(\pi) = ((p_1 * p_2) * p_3)(\pi)$.

Proof. Note that $(p_1 * (p_2 * p_3))(\pi) = (p_1 * p_{23})(\pi)$, and by definition,

$$(p_1 * p_{23})(\pi) = \sum_{\lambda \in S_N} (p_1(\pi) \circ \lambda^{-1}) p_{23}(\lambda).$$

Observe the following:

$$= \sum_{\lambda \in S_N} p_1(\pi \circ \lambda^{-1})(p_2 * p_3)(\lambda)$$

$$= \sum_{\lambda \in S_N} p_1(\pi \circ \lambda^{-1}) \sum_{\mu \in S_N} p_2(\lambda \circ \mu^{-1}) p_3(\mu)$$

$$= \sum_{\lambda \in S_N} (\sum_{\mu \in S_N} p_1(\pi \circ \lambda^{-1}) p_2(\lambda \circ \mu^{-1}) p_3(\mu)).$$

Then we let $\lambda = \sigma \circ \mu$ and $\sigma = \lambda \circ \mu^{-1}$.

$$= \sum_{\mu \in S_N} \left(\sum_{\sigma \in S_N} p_1(\pi \circ (\sigma \circ \mu)^{-1} p_2(\sigma) p_3(\mu)) \right)$$

$$= \sum_{\mu \in S_N} (\sum_{\sigma \in S_N} p_1((\pi \circ \mu^{-1}) \circ \sigma^{-1}) p_2(\sigma)) p_3(\mu)$$

$$=\sum_{\mu\in S_N} p_{12}(\pi\circ\mu^{-1})p_3(\mu)$$

$$=(p_{12}*p_3)(\pi)$$

$$= ((p_1 * p_2) * p_3)(\pi).$$

Thus we have that
$$(p_1 * (p_2 * p_3))(\pi) = ((p_1 * p_2) * p_3)(\pi)$$
.

For a deck of 52 cards to be completely randomized, we want each permutation on S_{52} to have probability $\frac{1}{52!}$. However, it is nearly impossible for this to occur. So when we ask how many shuffles are needed, we are not asking how far to go in order to achieve randomness, but rather how close to completely random is a shuffled deck? Next we are going to look at rising sequences within a deck, and we will find that if one is able to pick out which card has been inserted,

the deck has not been shuffled enough and is therefore not randomized.

5. RISING SEQUENCES

Definition 5.1. A **rising sequence** of a permutation is a maximal consecutively increasing subsequence.

How do we see rising sequences in a deck of cards?

The process is simple. First, we take a naturally ordered deck and perform a rearrangement from the permutations on S_N . We then pick any numbered card x and look after it to find card x + 1. If found, we then continue and look for x + 2. When we cannot find any further cards, we then repeat the process in reverse order. Now we start with card x - 1 and go backwards to find x - 2 and continue until no further cards can be found. This string of cards (...x-2, x-1, x, x+1, x+2, ...) is known as a rising sequence of a deck. We keep repeating this process until all cards of the deck are in a rising sequence.

Mann illustrates this concept well in [7] with an example: Say we have a deck of 8 cards, and we know that the order of these cards after a shuffle is 45162378. We may pick any card, but say we start with x = 3. We then look for card x + 1 = 4, and cannot find 4. So now we look before 3 and find 2 and 1. Thus, one of the rising sequences is 123. Next we start again with x = 6 and find 7 and 8 after it, and 4 and 5 before it. Thus the second rising sequence is 45678. We can now look at the order in this way:

$$45_{1}6_{23}78.$$

A trained eye may pick out rising sequences immediately. It is also noteworthy to notice that the order 45162378 can be obtained by cutting a naturally ordered deck after 3 cards and then performing a riffle shuffle.

Randomization is introduced here because the deck is not always split in half and it is not interlaced perfectly due to the fact that most people are clumsy shufflers.

When splitting a deck of N cards in half, the probability that k cards are chosen in first pile is given by the binomial distribution:

$$\binom{N}{k} 2^{-N}$$
.

After the deck has been cut into two piles, interlace the piles together

in anyway, so long as the cards from each pile maintain their relative order. This requirement comes natural to shufflers, because the cards on the bottom of each pile are dropped in order when shuffling.

Since there are $\binom{N}{k}$ possible interlacings, each possible interlacing has probability $1/\binom{N}{k}$. Thus, the probability of any given cut followed by any given interleaving is $\binom{N}{k}2^{-N}*1/\binom{N}{k}=1/2^N$

Next, the probability that the riffle shuffle produces the identity permutation (where all cards remain in fixed positions) is

$$\sum_{k=0}^{N} P(k \text{cards in first pile}) P(\text{identity} | k \text{cards in first pile})$$

$$= \sum_{k=0}^{N} {\binom{N}{k}} 2^{-N} \frac{1}{\binom{N}{k}} = (N+1)2^{-N}.$$

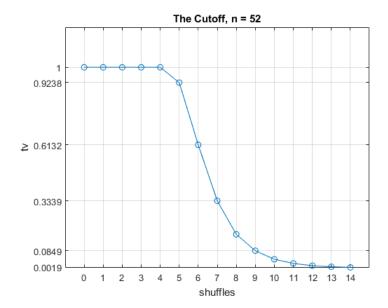
Thus, every permutation with one or two rising sequences other than identity has probability 2^{-N} of being chosen.

In [7], Mann gives the following example:

k = cut position	cut deck	probability of this cut	possible interleavings
0	123	1/8	123
1	1 23	3/8	123,213,231
2	12 3	3/8	123,132,312
3	123	1/8	123

permutation	[123]	[213]	[231]	[132]	[312]	[321]
probability under riffle shuffle	1/2	1/8	1/8	1/8	1/8	0

As a rule, a single riffle shuffle always gives a deck either 1 or 2 rising sequences. Being able to quickly pick out rising sequences in a deck is useful for magicians in performing some card tricks. Say that a spectator were to shuffle a new deck of cards 3 times, and then take the top card, look at it, and then reinsert in somewhere in the deck. Then a magician would observe anywhere from 3 to $2^3 = 8$ rising sequences and try to find the singleton rising sequence - which would be the reinserted card. Since the magician was able to pick out the singleton rising sequence, the deck has still not close to random after 3 shuffles. The question then remains, how many shuffles will bring a deck close to randomization?



Measuring how close a deck is to having a uniform distribution can be measured by the variation distance formula, where we are taking the difference between two probability distributions on S_N , Q_1 and Q_2 :

$$||Q_1 - Q_2|| = \frac{1}{2} \sum_{\pi \in S_n} |Q_1(\pi) - Q_2(\pi)|$$

The factor of 1/2 in the formula ensures that the number falls between 0 and 1.

When graphing the variation distance, the horizontal axis is the number of riffle shuffles (k), and the vertical axis is the variation distance to uniform. Note the variation distance for N=52 is near 1 for k=1,2,3,4 and then starts dropping rapidly. The graph makes a sharp cutoff at k=5, and by k=11 it has taken a value that is very near 0. We notice that a good middle point for the cutoff seems to be k=7.

By 7 shuffles, the possible number of rising sequences increases to 27 = 128. With their only being 52 cards in a deck, it would now be impossible to identify all rising sequences. Thus after 7 shuffles, our deck is very close to having a uniform distribution.

References

- [1] D. Aldous and P. Diaconis. "Shuffling cards and stopping times," *The Am. Math. Monthly* **93**, no. 5, (1986), 333-348.
- [2] W. Feller. An Introduction to Probability Theory and Its Applications, Vol. 1, 3rd Edition, Princeton University, 1950.
- [3] M. Gardner. Mathematical Magic Show. Knopf, New York, 1977.
- [4] M. Gardner. Mathematical Carnival. Knopf, New York, 1975.
- [5] G. Lawler and L Coyle. Lectures on Contemporary Probability, AMS, 1999.
- [6] P. Diaconis and R. Graham. Magical Mathematics, Prinston University Press, 2012.
- [7] B. Mann. "How many times should you shuffle a deck of cards?," in Snell, 1995.
- [8] A. Magid. "Interview with Martin Gardner," *Notices of the AMS*, Vol. 52, No. 6, 2005. http://www.ams.org/notices/200506/fea-gardner.pdf
- [9] J. O'Conner and E. Robertson. "Persi Warren Diaconis," 2011. http://www-history.mcs.st-andrews.ac.uk/Biographies/Diaconis.html
- [10] B. Alspach. "I'm in ...No, I'm out: Part 2," *Poker Digest* Vol. 2, No. 15, 16-29, 1999. http://people.math.sfu.ca/ alspach/mag26/

Cassidy Amason, Mathematics Department, Georgia College, Milledgeville, GA 31061, U.S.A.

E-mail address: cassidy.amason@bobcats.gcsu.edu