

**THE LINK BETWEEN SCRAMBLING NUMBERS
AND DERANGEMENTS**

Barry Balof, Eric Farmer, Jamie Kawabata

MS TR 97-02

May 1997

**Department of Mathematics
Rose-Hulman Institute of Technology
Terre Haute, IN 47803**

FAX (812) 877-3198

Phone: (812) 877-8391

The Link Between Scrambling Numbers and Derangements

Barry Balof, Eric Farmer, Jamie Kawabata *

May 6, 1997

Abstract

The group equation $abcdef = dabecf$ can be reduced to the equation $xcde = dxec$. In general, we are interested in how many variables are needed to represent group equations in which the right side is a permutation of the variables on the left side. Scrambling numbers capture this information about a permutation. In this paper we present several facts about scrambling numbers, and expose a striking relationship between permutations that cannot be reduced and derangements.

The group equation

$$abcdef = dabecf,$$

is associated with the permutation $(1, 2, 3, 5, 4)(6) \in S_6$ in a natural way:

$$abcdef = abcdef^{(1,2,3,5,4)(6)}.$$

In this case the group structure enables us to simplify the equation by cancelling f from both sides and replacing the product ab with its own symbol, say, $ab = x$. Thus the equation $abcdef = dabecf$, an equation in six variables, becomes

$$xcde = dxec = xcde^{(1,2,4,3)},$$

*Authors supported by NSF grant DMS-9322338

an equation in four variables with associated permutation $(1, 2, 4, 3) \in S_4$. In this example the equation cannot be simplified any further, i.e., the equation cannot be written using fewer than four variables.

Definition 1 *The scrambling number of a permutation $\pi \in S_n$, denoted $\text{scram}(\pi)$, is the smallest number of symbols needed to represent the n -variable group equation corresponding to π in the natural way.*

The scrambling number of the identity permutation, $\text{scram}(\text{id})$, is defined to be -1 as a matter of convention.

The number of permutations on n symbols with scrambling number k is denoted by $s_{n,k}$, i.e., $s_{n,k} = |\{\pi \in S_n \mid \text{scram}(\pi) = k\}|$. As a shorthand notation, we will write s_n for $s_{n,n}$.

We record the following facts concerning $s_{n,k}$:

$$\sum_{i=-1}^n s_{n,i} = n! \tag{1}$$

$$s_{n,-1} = 1 \tag{2}$$

$$s_{n,0} = s_{n,1} = 0 \tag{3}$$

$$s_{n,k} = \binom{n+1}{k+1} s_{k,k}, \quad 1 < k < n \tag{4}$$

Equations 1, 2, and 3 come directly from the definitions. The recursion formula for $s_{n,k}$ appearing in (4) was established by L. Smithline ([1]), and suggests that permutations on n symbols with scrambling number n are special.

Definition 2 *A permutation $\pi \in S_n$ is a perfect scrambling if $\text{scram}(\pi) = n$.*

THEOREM 1 $s_n = (n-2)s_{n-1} + 2(n-1)s_{n-2} + (n-1)s_{n-3}$, for $n \geq 3$.

Proof. Each perfect scrambling on n symbols can be constructed in one way by inserting an n th symbol, say f , into a permutation π on $n-1$ symbols. There are three cases we need to consider:

1. $\text{scram}(\pi) = n - 1$ (π is a perfect scrambling). The symbol f can be inserted in any of $n - 2$ positions. For example, in $baedc$, f can be inserted anywhere except at the right end (which would result in a cancellation) or immediately after the e (which would result in the two-symbol block ef). For this case the number of ways to construct perfect scramblings is $(n - 2)s_{n-1}$.
2. $\text{scram}(\pi) = n - 2$. There are three ways in which π can have scrambling number $n - 2$.
 - (a) If the first symbol cancels, as in $acedb$, then f must be inserted at the left end to prevent further cancellation. There are s_{n-2} permutations in which the first symbol cancels, and for each one we can only insert the n th symbol in one way.
 - (b) If two symbols act as a block, as in $daebc$, then f must split the block. There are $(n - 2)s_{n-2}$ permutations in which two symbols act as a block because there are $n - 2$ pairs that can act as a block and then s_{n-2} ways to permute the resulting $n - 2$ symbols, and for each such permutation we can only insert the n th symbol in one way.
 - (c) If the last symbol cancels, as in $badce$, then f may be inserted anywhere but at the right end. There are s_{n-2} permutations in which the last symbol cancels, and for each one the n th symbol can be inserted in $n - 1$ positions; anywhere except the right end.

In this case there are

$$s_{n-2} + (n - 2)s_{n-2} + (n - 1)s_{n-2} = 2(n - 1)s_{n-2}$$

ways to construct perfect scramblings.

3. $\text{scram}(\pi) = n - 3$. In this case, we cannot construct a perfect scrambling unless the last symbol cancels. For example, in $aecdb$ and $ecdab$ it is impossible to eliminate all cancellation and blocking with the insertion of f . If the last symbol does cancel, as in $\underline{a}dcbe$, $\underline{d}bcae$,

and $cbade$, the n th symbol can be inserted in one way (before the a , between the b and c , and between the d and e , respectively). There are $s_{n-2,n-3}$ permutations in which the last symbol cancels, and for each one we can insert the n th symbol in one way. From (4) we get $s_{n-2,n-3} = (n-1)s_{n-3}$ additional ways to construct perfect scramblings.

Considering these three cases is sufficient because if the scrambling number of π is less than $n-3$ there are multiple pairs of symbols together as blocks or multiple symbols cancelling at the ends, so it is impossible for the insertion of the n th symbol to result in a permutation without any blocking or cancellation.

The total for these three cases is

$$s_n = (n-2)s_{n-1} + 2(n-1)s_{n-2} + (n-1)s_{n-3}. \quad \square$$

We would like to find a closed form formula for s_n , and to do this we introduce an equivalent definition of $\text{scram}(\pi)$ in terms of two new objects that will be helpful in computing s_n .

Definition 3 For a given n , the set of all permutations $\pi \in S_n$ such that $\pi(i) + 1 = \pi(i+1)$ is called an **adjacency preserving set** and is denoted A_i .

Definition 4 For a given n , the set of all permutations $\pi \in S_n$ such that $\pi(i) = i$ is called a **point stabilizer** and is denoted F_i .

A permutation's corresponding equation can be reduced one symbol at a time by, at each step, grouping a pair of symbols into a block or cancelling a symbol. For example, the permutation $abefcd$ can be reduced to $abefx$, then $abyx$, then wyx , then yx . A grouping of two symbols into a block implies $\pi \in A_i$ for some i , and each cancellation implies $\pi \in F_1 \cup F_n$. For example, the reduction of $abefcd$ to $abefx$ implies $abefcd \in A_3$, the reduction of $abefx$ to $abyx$ implies $abefx \in A_5$,

and so on. The number of steps needed, k , is the number of elements of $\{F_1, A_1, A_2, \dots, A_{n-1}, F_n\}$ containing π , which is also the difference between n and the scrambling number of the permutation.

FACT 1 *The scrambling number of a permutation $\pi \in S_n$ is $n-k$, where k is the number of elements of $\{F_1, A_1, A_2, \dots, A_{n-1}, F_n\}$ containing π .*

This fact is our rationale for defining $\text{scram}(id)$ to be -1 . We also note that

1. A perfect scrambling $\pi \in S_n$ is a permutation that does not lie in any of the sets $F_1, A_1, A_2, \dots, A_{n-1}$, or F_n .
2. $|A_i| = (n-1)!$ for $1 \leq i \leq n-1$
3. $|F_i| = (n-1)!$ for $1 \leq i \leq n$
4. The cardinality of the intersection of k of the sets $F_1, A_1, A_2, \dots, A_{n-1}, F_n$ is $(n-k)!$.
5. The cardinality of the intersection of k of the sets $F_1, F_2, \dots, F_{n-1}, F_n$ is also $(n-k)!$.

These observations tell us that, among other things, the sets F_1 and F_n act just like the adjacency preserving sets. On this basis, we define $A_0 = F_1$ and $A_n = F_n$ to simplify our notation.

FACT 2

$$\begin{aligned}
 s_n &= n! - \binom{n+1}{1}(n-1)! + \binom{n+1}{2}(n-2)! - \dots + (-1)^{n+1} \\
 &= \sum_{k=0}^{n+1} (-1)^k \binom{n+1}{k} (n-k)! \tag{5}
 \end{aligned}$$

This fact is a straight forward application of the inclusion-exclusion principle to A_0, A_1, \dots, A_n .

Indeed, by recognizing that the derangements of n symbols are just permutations that do not lie in any of F_1, F_2, \dots, F_n , we have a striking similarity between (5) and the closed form for the

number of derangements, d_n , on n symbols:

$$\begin{aligned} d_n &= n! - \binom{n}{1}(n-1)! + \binom{n}{2}(n-2)! - \cdots + (-1)^n \\ &= \sum_{k=0}^n (-1)^k \binom{n}{k} (n-k)! \end{aligned} \tag{6}$$

THEOREM 2 $s_n + s_{n-1} = d_n$, for $n > 2$.

Proof. From Theorem 1 we know

$$s_n = (n-2)s_{n-1} + 2(n-1)s_{n-2} + (n-1)s_{n-3} \text{ for } n \geq 3,$$

and by rearranging, we get

$$s_n + s_{n-1} = (n-1)((s_{n-1} + s_{n-2}) + (s_{n-2} + s_{n-3}))$$

which looks like the well-known recursion relation for d_n ,

$$d_n = (n-1)(d_{n-1} + d_{n-2}).$$

If $s_k + s_{k-1} = d_k$ for all $k < n$ then we have

$$s_n + s_{n-1} = (n-1)(d_{n-1} + d_{n-2}) = d_n.$$

This, together with the fact that $s_2 + s_1 = d_2$ and $s_3 + s_2 = d_3$ gives us that $s_n + s_{n-1} = d_n$ for all $n > 2$ \square

Another Proof. To show that $s_n + s_{n-1} = d_n$, we replace each term with the proper closed-form formula to get

$$\sum_{k=0}^n (-1)^k \binom{n+1}{k} (n-k)! + \sum_{k=0}^{n-1} (-1)^k \binom{n}{k} (n-k-1)! = \sum_{k=0}^n (-1)^k \binom{n}{k} (n-k)!.$$

We then subtract $\sum_{k=0}^n (-1)^k \binom{n+1}{k} (n-k)!$ from both sides of the equation and manipulate the equation to an identity.

$$\begin{aligned}
\sum_{k=0}^{n-1} (-1)^k \binom{n}{k} (n-k-1)! &= \sum_{k=0}^n (-1)^k \binom{n}{k} (n-k)! - \sum_{k=0}^n (-1)^k \binom{n+1}{k} (n-k)! \\
&= \sum_{k=0}^n (-1)^k \left(\binom{n}{k} - \binom{n+1}{k} \right) (n-k)! \\
&= \sum_{k=0}^n (-1)^k \left(\binom{n}{k-1} (-1) \right) (n-k)! \\
&= \sum_{k=-1}^{n-1} (-1)^{k+1} \left(\binom{n}{k} (-1) \right) (n-k-1)! \\
&= \sum_{k=-1}^{n-1} (-1)^k \binom{n}{k} (n-k-1)! \\
&= \sum_{k=0}^{n-1} (-1)^k \binom{n}{k} (n-k-1)! \quad \square
\end{aligned}$$

With the machinery we have developed so far (the recurrence relation and closed form formula) we know a lot about the distribution of $s_{n,k}$, which is in several ways similar to the distribution of the number of fixed points in a permutation.

THEOREM 3 $\lim_{n \rightarrow \infty} \frac{s_n}{n!} = \frac{1}{e}$

Proof. Obviously $s_{n-1} \leq (n-1)!$, so $\lim_{n \rightarrow \infty} \frac{s_{n-1}}{n!} = 0$, and

$$\lim_{n \rightarrow \infty} \frac{s_n}{n!} = \lim_{n \rightarrow \infty} \frac{d_n}{n!} - \lim_{n \rightarrow \infty} \frac{s_{n-1}}{n!} = \frac{1}{e}.$$

THEOREM 4 For all n , the mean of $\text{scram}(\pi)$ over S_n is $n - 1 - \frac{1}{n}$.

Proof. Define $\chi : S_n \times \{0, 1, \dots, n\} \rightarrow \{0, 1\}$ by:

$$\chi(\pi, i) = \begin{cases} 1 & \text{if } \pi \in A_i \\ 0 & \text{otherwise} \end{cases} \quad \text{for } 0 \leq i \leq n$$

Note that for all $\pi \in S_n$, $\text{scram}(\pi) = n - \sum_{i=0}^n \chi(\pi, i)$, so

$$\begin{aligned}
E(\text{scram}(\pi)) &= \frac{1}{|S_n|} \sum_{\pi \in S_n} \left(n - \sum_{i=0}^n \chi(\pi, i) \right) \\
&= \frac{1}{n!} \left(n \cdot n! - \sum_{\pi \in S_n} \sum_{i=0}^n \chi(\pi, i) \right) \\
&= n - \frac{1}{n!} \sum_{\pi \in S_n} \sum_{i=0}^n \chi(\pi, i) \\
&= n - \frac{1}{n!} \sum_{i=0}^n \sum_{\pi \in S_n} \chi(\pi, i)
\end{aligned}$$

and since $\sum_{\pi \in S_n} \chi(\pi, i) = |A_i| = (n-1)!$ for $0 \leq i \leq n$, we have

$$\begin{aligned}
E(\text{scram}(\pi)) &= n - \frac{1}{n!} \sum_{i=0}^n (n-1)! \\
&= n - \frac{(n+1)(n-1)!}{n!} \\
&= n - 1 - \frac{1}{n} \quad \square
\end{aligned}$$

THEOREM 5 For all n , the variance of $\text{scram}(\pi)$ over all $\pi \in S_n$ is $\frac{n+1}{n-1} - \frac{1}{n} - \frac{1}{n^2}$

Proof. The variance, σ^2 is the difference between the mean of the squares and the square of the mean.

$$\sigma^2 = E(\text{scram}(\pi)^2) - E(\text{scram}(\pi))^2$$

The square of the mean, $E(\text{scram}(\pi))^2$, can be found easily from theorem 4. The mean of the squares, $E(\text{scram}(\pi)^2)$, can be computed as follows.

$$\begin{aligned}
E(\text{scram}(\pi)^2) &= \frac{1}{n!} \sum_{\pi \in S_n} \text{scram}(\pi)^2 \\
&= \frac{1}{n!} \sum_{\pi \in S_n} \left(n - \sum_{i=0}^n \chi(\pi, i) \right)^2 \\
&= \frac{1}{n!} \sum_{\pi \in S_n} \left(n^2 - 2n \sum_{i=0}^n \chi(\pi, i) + \left(\sum_{i=0}^n \chi(\pi, i) \right)^2 \right)
\end{aligned}$$

$$\begin{aligned}
&= \frac{1}{n!} \sum_{\pi \in S_n} \left(n^2 - 2n \sum_{i=0}^n \chi(\pi, i) + \sum_{i=0}^n \sum_{j=0}^n \chi(\pi, i) \chi(\pi, j) \right) \\
&= \frac{1}{n!} \left(\sum_{\pi \in S_n} n^2 - 2n \sum_{i=0}^n \sum_{\pi \in S_n} \chi(\pi, i) + \sum_{i=0}^n \sum_{j=0}^n \sum_{\pi \in S_n} \chi(\pi, i) \chi(\pi, j) \right) \\
&= \frac{1}{n!} \left(n!n^2 - 2n(n+1)(n-1)! + \sum_{i=0}^n \sum_{j=0}^n \sum_{\pi \in S_n} \chi(\pi, i) \chi(\pi, j) \right) \\
&= \frac{1}{n!} \left(n!n^2 - 2(n+1)! + \sum_{i=0}^n \left(\sum_{j \neq i} \sum_{\pi \in S_n} \chi(\pi, i) \chi(\pi, j) + \sum_{\pi \in S_n} \chi(\pi, i) \chi(\pi, i) \right) \right) \\
&= n^2 - 2(n+1) + \frac{1}{n!} \sum_{i=0}^n \left(\sum_{j \neq i} (n-2)! + (n-1)! \right) \\
&= n^2 - 2(n+1) + \frac{1}{n!} (n+1)(n(n-2)! + (n-1)!) \\
&= n^2 - 2n - 2 + \frac{n+1}{n-1} + \frac{n+1}{n} \\
&= n^2 - 2n - 1 + \frac{n+1}{n-1} + \frac{1}{n}
\end{aligned}$$

From Theorem 4 we have

$$E(\text{scram}(\pi))^2 = (n-1 - 1/n)^2 = n^2 - 2n - 1 + \frac{2}{n} + \frac{1}{n^2}.$$

Putting the two together, we get

$$\begin{aligned}
\sigma^2 &= \left(n^2 - 2n - 1 + \frac{n+1}{n-1} + \frac{1}{n} \right) - \left(n^2 - 2n - 1 + \frac{2}{n} + \frac{1}{n^2} \right) \\
&= \frac{n+1}{n-1} - \frac{1}{n} - \frac{1}{n^2} \quad \square
\end{aligned}$$

These strong numerical relationships lead us to believe that there is a relatively simple bijection argument relating the two. The task of finding this bijection would be simpler if we were trying to find a bijection from the derangements to a single structure of the same size rather than to the perfect scramblings on n and $n-1$ symbols. To accomplish this we introduce two new kinds of scramblings which have nice combinatorial properties.

Definition 5 A secondary scrambling is an element of S_n in which there are no adjacencies and the left endpoint is not fixed.

The number of secondary scramblings in S_n is denoted s'_n , i.e., $s'_n = |S_n - A_0 \cup A_1 \cup \dots \cup A_{n-1}|$.

These permutations may or may not lie in A_n .

Definition 6 A tertiary scrambling is an element of S_n in which there are no adjacencies, though both endpoints may or may not be fixed.

The number of tertiary scramblings in S_n is denoted s''_n , i.e., $s''_n = |S_n - A_1 \cup A_2 \cup \dots \cup A_{n-1}|$.

These permutations may lie in A_0 or A_n or both.

THEOREM 6 $s'_n = s_n + s_{n-1}$ and $s''_n = s'_n + s'_{n-1}$.

Proof. One can easily construct all secondary scramblings on n symbols from perfect scramblings on n and $n - 1$ symbols. All perfect scramblings on n symbols are already near-perfect scramblings on n symbols. The rest of the near-perfect scramblings can be gotten by appending an n -th symbol to a perfect scrambling on $n - 1$ symbols.

A similar construction will construct all tertiary scramblings on n symbols from secondary scramblings on n and $n - 1$ symbols. The secondary scramblings on n symbols are also tertiary scramblings on n symbols, and the remaining tertiary scramblings can be constructed by prepending a zero-th symbol to a near-perfect scramblings on $n - 1$ symbols and renaming all the symbols so they range from 1 to n rather than from 0 to $n - 1$. \square

With this theorem and Theorem 2 we can look for any of three forms of the bijection.

$$s_n + s_{n-1} = d_n$$

$$s'_n = d_n$$

$$s''_n = d_n + d_{n-1}$$

A bijective proof of any of these three equations would give us a bijective proof of the other two, since Theorem 6 uses a bijective argument and the composition of two bijections is a bijection. Of particular interest is the second equation, $s'_n = d_n$, since the bijection we are looking for is one between two objects of similar structure:

$$s'_n = |S_n - A_1 \cup A_2 \cup \dots \cup A_n|$$

$$d_n = |S_n - F_1 \cup F_2 \cup \dots \cup F_n|$$

We will see that indeed, the structure of derangements and secondary scramblings are very similar, and we will be able to use this structural similarity to find a bijection between them.

We are going to look at the combinatorics behind the recursion relation $d_n = (n-1)(d_{n-1} + d_{n-2})$ to find steps to build any derangement from a derangement on fewer symbols. We will then do the same thing with the recursion relation $s'_n = (n-1)(s_{n-1} + s_{n-2})$, finding steps to build any near-perfect scrambling from near-perfect scramblings on fewer symbols. We will see that there is a simple correspondence between the derangement construction steps and near-perfect scrambling construction steps. Our bijection will then consist of decomposing a derangement into a (unique) sequence of derangement construction steps, translating these steps into a sequence of near-perfect scrambling construction steps, and then executing the steps to build the corresponding near-perfect scrambling.

First we consider the combinatorial argument behind the recursion relation $d_n = (n-1)(d_{n-1} + d_{n-2})$. We can build a derangement on n symbols from a derangement on $n-1$ symbols by appending an n -th symbol and then swapping it with any of the other $n-1$ symbols. This will yield $(n-1)d_{n-1}$ derangements. The rest of the derangements can be constructed from permutations on $n-1$ symbols with one fixed point by simply appending an n -th symbol and swapping it with the other fixed point. There are exactly $(n-1)d_{n-2}$ permutations on $n-1$ symbols with one fixed point, and any permutation with more than one fixed point cannot be used in this way to build a

derangement.

This gives us the derangement-construction steps we are looking for. We define a function α_i to take a derangement on $n - 1$ symbols, append an n -th symbol, and then swap it with the i -th symbol. So, for example, $\alpha_2(badc) = bedca$. We can also build a derangement on n symbols from a permutation on $n - 1$ symbols with one fixed point, and the permutation on $n - 1$ symbols with one fixed point can be in turn built from a derangement on $n - 2$ fixed points by fixing the i -th point and applying the derangement on $n - 2$ symbols to the rest. We define the function β_i to take a derangement on $n - 2$ symbols, build a permutation on $n - 1$ symbols with the i -th point fixed, and then building from that a derangement on n symbols. So, for example, $\beta_3(badc) = bafedc$, with $baced$ as the intermediate permutation with one fixed point.

Now we seek to find a similar combinatorial argument behind $s'_n = (n - 1)(s'_{n-1} + s'_{n-2})$. We can build a near-perfect scrambling on n symbols by inserting an n -th symbol into a near-perfect scrambling on $n - 1$ symbols. We can insert the n -th symbol anywhere (including the right end) as long as we do not insert it immediately to the right of the $n - 1$ -st symbol. This gives us $n - 1$ places to insert, and a total of $(n - 1)s'_{n-1}$ near-perfect scramblings that can be built in this way. We can also build near-perfect scramblings from permutations that are not near-perfect scramblings. If the first symbol cancels, or if two symbols are together in a block, we can insert the n -th symbol before the first symbol to prevent cancellation, or between the two symbols to break up the block. It is not hard to see that there are $(n - 1)s'_{n-2}$ permutations of this kind. Each can be built by taking a near-perfect scrambling on $n - 2$ symbols and "unreducing" once, where unreducing consists of inserting a fixed point at the right end or expanding any of the $n - 2$ symbols into a block of two symbols. This gives us $n - 1$ ways to unreduce any near-perfect scrambling on $n - 2$ symbols, accounting for the $(n - 1)s'_{n-2}$ term in the recursion relation.

With this reasoning behind the recursion relation, it is easy to see what the near-perfect scram-

bling construction steps are going to be. We let the function γ_i take a near-perfect scrambling on $n - 1$ symbols and insert an n -th symbol in the i -th legal spot, counting from left to right. So, for example, $\gamma_3(bdac) = bdaec$. Note that the third *legal* insertion point is the fourth insertion point because e cannot be inserted after d . We can also build a near-perfect scrambling on n symbols from a near-perfect scrambling on $n - 2$ symbols by unreducing in one of $n - 1$ ways and then inserting the n -th symbol in the single legal spot. We define the function δ_i to do just this. δ_i takes a near-perfect scrambling on $n - 2$ symbols, unreduces it to a permutation on $n - 1$ symbols, and then inserts the n -th symbol, where the unreduction depends on i . For $i = 1$, a fixed point is inserted at the left end, and for $i > 1$, the $i - 1$ -st symbol (from left to right) is expanded into a block of two symbols. So, for example $\delta_1(bdac) = facebd$, with $acebd$ as the intermediate unreduced permutation. To see another example, $\delta_3(bdac) = bdfaec$, with $bdeac$ as the intermediate unreduced permutation.

By observing the natural correspondence between α and γ and between β and δ , we have the bijection we are looking for. Simply take a derangement, such as $ecdab$, decompose it into its unique sequence of steps, in this case $\alpha_1(\alpha_3(\alpha_2(ba)))$, translate the steps into the corresponding steps for near-perfect scramblings, which are $\gamma_1(\gamma_3(\gamma_2(ba)))$, and then execute the near-perfect scrambling construction steps, which in this case yield the near-perfect scrambling $ebadc$. The other direction of the bijection works in exactly the same way.

References

- [1] L. SMITHLINE Rewritability, Commutators, and Fundamental n -rewritings.