



University of Pennsylvania  
ScholarlyCommons

---

Finance Papers

Wharton Faculty Research

---

6-2016

# Buses, Bullies, and Bijections

Vladimir Pozdnyakov

J. Michael Steele  
*University of Pennsylvania*

Follow this and additional works at: [http://repository.upenn.edu/fnce\\_papers](http://repository.upenn.edu/fnce_papers)

 Part of the [Finance and Financial Management Commons](#)

---

## Recommended Citation

Pozdnyakov, V., & Steele, J. (2016). Buses, Bullies, and Bijections. *Mathematics Magazine*, 89 (3), 167-176. <http://dx.doi.org/10.4169/math.mag.89.3.167>

This paper is posted at ScholarlyCommons. [http://repository.upenn.edu/fnce\\_papers/151](http://repository.upenn.edu/fnce_papers/151)  
For more information, please contact [repository@pobox.upenn.edu](mailto:repository@pobox.upenn.edu).

---

# Buses, Bullies, and Bijections

## **Abstract**

The random—or orderly—seating of passengers on a bus is used to motivate several questions about cycles of permutations. These in turn motivates the investigation of bijections between special subsets of permutations. The goal, of course, is to give simple explanations of surprising facts.

## **Disciplines**

Finance and Financial Management

# Buses, Bullies, and Bijections

VLADIMIR POZDNYAKOV  
 University of Connecticut  
 Storrs, CT 06269  
[vladimir.pozdnyakov@uconn.edu](mailto:vladimir.pozdnyakov@uconn.edu)

J. MICHAEL STEELE  
 University of Pennsylvania  
 Philadelphia, PA 19104  
[steele@wharton.upenn.edu](mailto:steele@wharton.upenn.edu)

## A bully comes onto a bus

In times past in a country far away, passengers on intercity buses were all assigned seat numbers, and the buses were always full. There were many casual people, and there were a few people who were real sticklers for the rules. This is where our problem begins.

Adam is the first person to enter an  $n$  seat intercity bus, and, since he is a casual person, he takes a seat at random. Subsequently,  $n - 2$  more casual people enter the bus, and they also happily take seats at random. Of course, the inevitable occurs, and the last person on the bus is one of the aforementioned sticklers for the rules.

If luck prevails, the stickler's assigned seat is free; he sits down, and the bus drives away. On the other hand, if his seat is occupied, the stickler insists on having his assigned seat, and he even insists that the person whom he displaces must go to his or her officially assigned seat. Moreover, the stickler continues to oversee the commotion he has created. He insists that each successively displaced person must go to his or her assigned seat until there are no more displaced persons. The  $n - 1$  casual people are not happy about this; but, eventually, the brouhaha dies down, and the bus hits the road.

Here is the question: "What is the probability that mellow Adam, the first person on the bus, will be forced to move from his randomly chosen seat?"

If  $n = 2$ , it is easy to see that the answer is  $1/2$ . If Adam happens to sit in his assigned seat, then everything is fine; otherwise, he will have to move. Next, we can consider a bus with  $n = 3$  seats, but things become more complicated. For a three seat bus, there are  $6 = 3!$  possibilities that one needs to consider. To work through these possibilities in a systematic way, one needs some tools like the notation that we develop in the next section. Still, after an examination of the six cases, one comes to a noteworthy observation. In exactly half of these cases, Adam is forced to move.

This coincidence suggests a bold speculation: "Can it be true that the probability that Adam will be forced to move is always one-half whatever the size of the bus may be?"

## Representing a permutation in two or more ways

In purely mathematical terms, we have a problem about a randomly chosen permutation of the set  $[n] = \{1, 2, \dots, n\}$ , whereby a *permutation* of  $[n]$  we just mean a bijection (or one-to-one correspondence) between the elements of the set  $[n]$ . The prolific

French mathematician Augustin-Louis Cauchy (1789–1857) suggested that to describe a permutation one should write the elements of  $[n]$  in their natural order in one row and then write the values to which they are mapped in a second row. For example, to describe a typical permutation of  $[6] = \{1, 2, \dots, 6\}$  one might write

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 3 & 5 & 6 & 1 & 4 \end{pmatrix}, \quad (1)$$

by which one would mean that 1 is mapped to 2, and 2 is mapped to 3, and 3 is mapped to 5, and so on. This representation is called the *two-line notation* for the permutation  $\sigma$ , and, after writing a few such representations, one realizes that in most situations the first line can be safely omitted. When one represents a permutation  $\sigma$  with just the second line, we have what is unsurprisingly called the *one-line notation* for a permutation.

There is a third notation for permutations that offer some serious benefits, but its construction calls for a little more thought. The underlying idea is that one takes a value from the set  $[n]$ , and one then examines where that value is mapped by successive applications of the function  $\sigma : [n] \rightarrow [n]$ . For example, if we take the  $\sigma$  given by the two-lined formula (1) and if we start with 1, then we find the successive images

$$1 \mapsto 2 \mapsto 3 \mapsto 5 \mapsto 1. \quad (2)$$

We could record this information by putting all of the indicated values between parentheses, but that would be a bit redundant since 1 would appear twice. We lose no information if we clip off one of the 1's, and for the moment we will just clip off the last 1. We then record our four element cycle as  $(1, 2, 3, 5)$ .

We then look at the smallest number that is not in this cycle; in this example that would be 4. To find the cycle containing 4 we again follow the assignments given by Equation (1), and we find  $4 \mapsto 6 \mapsto 4$ . As before, we drop the repeated value at the end, and we record this cycle information by writing  $(4, 6)$ . After we have found all of the cycles, we then need to choose an order in which to present them. One natural idea is to order the cycles so that their smallest elements are increasing. In the end, this gives us what we will call the *first cycle representation* of  $\sigma$ ,

$$\sigma = (1, 2, 3, 5)(4, 6).$$

## Back to the brouhaha on the bus

Returning to the bus, we label the passengers 1 through  $n$ , and we assume for each  $k \in [n]$  that passenger  $k$  is officially assigned seat  $k$ . Also, without loss of generality, we assume that the stickler is passenger number 1 and that Adam is passenger number 2. Next, we consider a random permutation  $\sigma : [n] \rightarrow [n]$ , and for  $2 \leq k \leq n$  we let  $\sigma(k)$  denote the seat that that passenger  $k$  has randomly occupied. By elimination, seat  $\sigma(1)$  is the only seat that is empty when the stickler boards the bus.

If  $\sigma(1) = 1$  the stickler sits in seat 1, and the bus departs. In such a case, one should note that in the random permutation  $\sigma$ , we have that 1 is in a cycle of length 1. On the other hand, if  $\sigma(1) \neq 1$ , then the stickler refuses to sit in the empty seat  $\sigma(1)$ ; he wants his assigned seat number 1. This then begins a chain of grumbling dislocations.

To see how the dislocations evolve in a concrete example, we can reuse the sample permutation  $\sigma$  that we used in Equation (1) to illustrate Cauchy's notation. Here we identify an element  $i$  of the top row with person  $i$ , and, for  $2 \leq i \leq 6$ , we let the values in the second row denote the seat numbers that have been occupied at random by the corresponding persons from the top row.

Just before the stickler arrives, only seat 2 is empty; but, when the stickler arrives, he ignores the empty seat. Instead, marches directly to seat 1 where person 5 is sitting. He then makes person 5 move to seat 5. Unfortunately seat 5 is occupied by person 3, so person 3 has to move to seat 3. This seat is occupied by person 2 who then has to move to seat 2. Thankfully seat 2 is empty, so the shuffling can finally end.

The net effect of this drama is that the individuals  $\{2, 3, 5\}$  are the ones who had to move, and the key observation is that these values and  $\{1\}$  are exactly the values that appear in the first cycle of the cycle representation of  $\sigma$ . It's not hard to see that this example is generic; no matter what the size of  $n$  or what the random permutation  $\sigma$ , Adam will have to move if and only if Adam is in the first cycle of the cycle representation of  $\sigma$ .

Given this observation, we have a natural plan for finding the probability of Adam having to move. We just need to solve a purely combinatorial problem: "How many permutations of  $[n]$  have 1 and 2 in the same cycle?"

### A classical counting calculation: Possibly a straw man

Such a natural question is bound to have been asked before, and in fact it has been posed at least once as an exercise (cf. [9] p. 58). We will see shortly that there is a lovely bijection that makes it completely obvious that the answer is  $n!/2$ . Still, why not first give a nod to Kipling and the *Gods of the Copy Book Headings*? In other words, why not be humble and simply do our sums?

Let  $k$  be the size of a cycle that contains both 1 and 2. We have  $2 \leq k \leq n$ , and there are  $\binom{n-2}{k-2}$  ways to select the other elements of  $[n] \setminus \{1, 2\}$  that are in the cycle of size  $k$  that contains 1 and 2. Given the elements for this cycle of size  $k$ , there are then  $(k-1)!$  inequivalent ways to order this cycle. Finally, there are  $n-k$  elements of  $[n]$  remaining, and these can be ordered in  $(n-k)!$  ways to make a permutation. By following the first cycle with this permutation, we have defined a one-to-one mapping from  $[n]$  onto  $[n]$ , i.e., we have specified a complete permutation of  $[n]$ .

Now we do our sums. Specifically, we let  $Q$  denote the number of permutations that have 1 and 2 in the same cycle, so, when we condition on  $k \in \{2, 3, \dots, n\}$ , our earlier observations give us

$$\begin{aligned} Q &= \sum_{k=2}^n \binom{n-2}{k-2} (k-1)!(n-k)! = \sum_{k=2}^n \frac{(n-2)!}{(n-k)!(k-2)!} (k-1)!(n-k)! \\ &= (n-2)! \sum_{k=2}^n (k-1) = (n-2)! \sum_{k=1}^{n-1} k = (n-2)! \frac{(n-1)n}{2} = \frac{n!}{2}. \end{aligned}$$

The charm of this calculation is that it offers good practice with some basic tools of combinatorial counting—plus, of course, it *does* get the job done. Now we know for sure that Adam has to move with probability  $1/2$ —no matter what the size of the bus may be.

On the other hand, just because a problem is solved, it does not mean that one cannot gain more insight. In fact, as promised earlier, there is a simple bijection that recovers the relationships  $Q = n!/2$  without the need for any real calculation. Moreover, the same bijection also gives quick answers to more complex problems.

## Want insight? Consider a bijection!

The classical calculation got us the answer we hoped to find, but the calculations did not really make us *see* why Adam always has to move with probability one-half. We can do better with a bijection.

As usual, we let  $\mathbb{S}_n$  denote the set of permutations of  $[n]$ , and now we consider a partition of  $\mathbb{S}_n$  into two subsets. Specifically, we let  $\mathcal{A}_{(1)(2)}$  be the subset of elements of  $\mathbb{S}_n$  that have 1 and 2 in different cycles, and we let  $\mathcal{A}_{(1,2)}$  be the subset of elements of  $\mathbb{S}_n$  that have 1 and 2 in the same cycle. In terms of the cycle representation we can write

$$\begin{aligned}\mathcal{A}_{(1)(2)} &= \{\sigma \in \mathbb{S}_n : (1 \dots)(2 \dots) \dots\} \quad \text{and} \\ \mathcal{A}_{(1,2)} &= \{\sigma \in \mathbb{S}_n : (1 \dots 2 \dots) \dots\}.\end{aligned}$$

In each of these formulas the first two ellipses (i.e., the  $\dots$ 's) can then be any pair of disjoint subsets of  $[n] \setminus \{1, 2\}$ , including possibly the empty set.

Now consider the transformation  $U : \mathcal{A}_{(1)(2)} \rightarrow \mathcal{A}_{(1,2)}$  that one gets by erasing the back-to-back parenthesis pair “)(” that precedes the 2 in a permutation  $\sigma \in \mathcal{A}_{(1)(2)}$ . Thus, for example, if  $n = 7$ , then one would have

$$\sigma = (153)(24)(67) \mapsto U(\sigma) = (15324)(67) = \tau.$$

To define the inverse transformation  $U^{-1} : \mathcal{A}_{(1,2)} \rightarrow \mathcal{A}_{(1)(2)}$  we just reverse our recipe; that is, we insert the parenthesis grouping in front of the 2 in a given  $\tau \in \mathcal{A}_{(1,2)}$ . To continue with our example, one would have

$$\tau = (15324)(67) \mapsto U^{-1}(\tau) = (153)(24)(67) = \sigma.$$

Since we have a bijection between  $\mathcal{A}_{(1)(2)}$  and  $\mathcal{A}_{(1,2)}$  the two sets have equal cardinalities. Moreover, since their disjoint union is  $\mathbb{S}_n$ , the sum of these cardinalities is  $n!$  and we have

$$|\mathcal{A}_{(1)(2)}| = |\mathcal{A}_{(1,2)}| = \frac{1}{2}n!.$$

This formula seems to provide a richer understanding of the reason why Adam has to move with probability  $1/2$ . It's because there is a bijection between  $\mathcal{A}_{(1)(2)}$  and  $\mathcal{A}_{(1,2)}$ . Moreover, the bijection is very simple: we just erase the first pair of back-to-back parentheses in the cycle representation of the permutation.

## Building a better bijection from $\mathbb{S}_n$ onto $\mathbb{S}_n$

In our continuing example, we began with a permutation  $\sigma$  with Cauchy's representation (1), and we found that there were two cycles that one could write as paths with the same head and tail:

$$1 \mapsto 2 \mapsto 3 \mapsto 5 \mapsto 1 \quad \text{and} \quad 4 \mapsto 6 \mapsto 4. \quad (3)$$

We then built an efficient cycle representation of  $\sigma$  by cutting of the tails of these paths and ordering the resulting strings in increasing order of their heads. The benefit of this representation became evident when it lead us to the surprisingly simple bijection between  $\mathcal{A}_{(1)(2)}$  and  $\mathcal{A}_{(1,2)}$ .

Still, when two paths diverge in a wood, it sometimes pays to go back to see where the other path may have led. Let's consider cutting off the heads in (3) and keeping the

tails. We can then order the resulting strings in increasing order of the tails, and we have a *second recipe* for going from a permutation to a uniquely string of parentheses and numbers. For example if  $\sigma = [3, 6, 5, 8, 1, 2, 4, 7]$  in one-line notation, then our two recipes give us the strings:

$$\begin{aligned} \text{First Recipe: } \sigma &= (1, 3, 5)(2, 6)(4, 8, 7), \\ \text{Second Recipe: } \sigma &= (3, 5, 1)(6, 2)(8, 7, 4). \end{aligned} \quad (4)$$

When viewed as products of cycles, each of these strings represent the same permutation. Nevertheless, if we ignore the interpretation of these formulas and just look at them as the strings of symbols, then some interesting distinctions emerge. For example, consider a new erasure operation where we erase *all* of the back-to-back parentheses. We then interpret the resulting string as a permutation in one-line notation.

Each of the two recipes then gives us mapping from  $\mathbb{S}_n$  to  $\mathbb{S}_n$ . If we mnemonically call the first mapping  $F$  and the second mapping  $S$ , then for our example we have

$$\begin{aligned} \text{Using the First Recipe: } F(\sigma) &= [1, 3, 5, 2, 6, 4, 8, 7], \\ \text{Using the Second Recipe: } S(\sigma) &= [3, 5, 1, 6, 2, 8, 7, 4]. \end{aligned}$$

Curiously enough, the mapping  $F : \mathbb{S}_n \rightarrow \mathbb{S}_n$  has a serious limitation. It is not surjective. To see the problem, just note that for any  $\sigma \in \mathbb{S}_n$  the first element of the string  $F(\sigma)$  is always equal to 1, so it cannot be a surjection. To be completely explicit just note that  $\tau = [2, 3, 4, 5, 6, 7, 1]$  is not equal to  $F(\sigma)$  for any  $\sigma$ .

On the other hand, the mapping  $S : \mathbb{S}_n \rightarrow \mathbb{S}_n$  is an honest bijection—and a stunningly useful one to boot! Since  $S$  maps the *finite* set  $\mathbb{S}_n$  into itself, the map  $S$  must be injective if it is surjective. Thus, to show that  $S$  is a bijection, we just need to show that for each  $\tau \in \mathbb{S}_n$  there is a  $\sigma \in \mathbb{S}_n$  such that  $S(\sigma) = \tau$ .

We will show this by means of an algorithm. Specifically, we take  $\tau \in \mathbb{S}_n$  in its one-line form  $[a_1, a_2, \dots, a_n]$ , and we follow a four step process to find a  $\sigma$  such that  $S(\sigma) = \tau$ .

- Step 1:** Scan  $\tau$  until we come to 1, then put down back-to-back parentheses after 1, unless we are at the end—in which case we just stop.
- Step 2:** Compute the smallest value  $s$  that has not yet been scanned.
- Step 3:** Continue scanning the rest of  $\tau$  until arriving at  $s$ , then put down back-to-back parentheses after  $s$ , unless we are at the end—in which case we just stop.
- Step 4:** Repeat Steps 2 and 3 until done.

To see how the algorithm works, one can take  $\tau = [7, 1, 4, 5, 2, 3, 6]$ . After Steps 1 and 2, we have the intermediate result  $(7, 1)(4, 5, 2, 3, 6)$  and  $s = 2$ . When we apply Steps 3 and 2 we have  $(7, 1)(4, 5, 2), (3, 6)$  and  $s = 3$ . Finally, after another repetition of Step 3, we have  $\sigma = (7, 1)(4, 5, 2)(3)(6)$ , and we can safely stop. It is trivial to check that we have the correct result; when we apply the process  $S$  of internal parenthesis removal on  $\sigma$ , we have  $S(\sigma) = \tau$ . Incidentally, this is a good time to note that a cycle of size one is the same as a fixed point, e.g., if 3 is a fixed point of  $\sigma$ , then  $\sigma(3) = 3$  and the cycle of  $\sigma$  that contains 3 is simply (3).

At the end of the day we have confirmed that  $S : \mathbb{S}_n \mapsto \mathbb{S}_n$  is an honest bijection, but we face some natural questions. Why is this interesting? What can one do with the bijection? We can answer these questions by quizzing ourselves a little harder about Adam and the other boarders of the bus.

## Back on the bus with new questions

Suppose that Adam boarded the bus at the same time as four other people from work. What is the probability that the stickler's insistence on the rules will force *all* five of these workers to move from their seats?

In more mathematical terms, we take  $2 \leq k \leq n$ , and we ask for the probability that all of the values  $\{2, 3, \dots, k\}$  happen to be in the first cycle of a random permutation of  $[n]$ . Remarkably, this probability again turns out not to depend on  $n$ , and one can see this with help from the bijections  $S$  and  $S^{-1}$ .

Suppose that  $\tau = [a_1, a_2, \dots, a_n]$  is a random ordering of  $[n]$  that we view as the one-line notation of a random permutation  $\tau$ . One can generate such a random permutation by sampling  $n$  times without replacement from the set  $[n]$ . By the one-to-one correspondence given by  $S^{-1}$  this permutation is paired with a unique permutation  $\sigma$  that is given in cycle notation and whose first cycle is  $(a_1, a_2, \dots, a_j)$ , where  $a_j = 1$  and  $1 \leq j \leq n$ .

What's the bottom line? We now see that the probability that the values  $\{2, 3, \dots, k\}$  are all in the same cycle of  $\sigma$  as 1 is the same as the probability that all of the values  $\{2, 3, \dots, k\}$  precede 1 in the one-line permutation  $\tau = [a_1, a_2, \dots, a_n]$ . In a random ordering of  $[n]$  the values of  $[k]$  also appear in random order, so the probability that  $\{2, 3, \dots, k\}$  precede 1 in  $\tau$  is the same as the probability that 1 is the last number in a random ordering of  $[k]$ . In such a random ordering, every placement of 1 is equally likely, so the probability that 1 comes last is  $1/k$ .

Thus, the probability that the values  $\{2, 3, \dots, k\}$  are all in the first cycle of a random permutation of  $[n]$  is exactly  $1/k$  for all  $2 \leq k \leq n$ . For  $k = 2$  we get  $1/2$ , and this is just the result we found from our first bijective argument. For  $k = n$ , this just reflects that the number of permutations of  $[n]$  with 1 in the last position is just  $1/n$ . To fill in all the values in between, we just use our bijection!

Parenthetically, we should note that the bijection constructed here is essentially the same as a standard ordering that is widely used (e.g., **[6]** or **[1]**). The only distinction is that the more common standard ordering breaks  $\tau$  at the successive maxima, and, here, for compatibility with our first bijection, we have broken  $\tau$  at the successive minima.

## How many get bumped?

If the set of co-workers on the bus is identified with the set  $\{2, 3, \dots, k\}$  where we have  $2 \leq k \leq n$ , then it is also natural to ask about the *total number* of co-workers who get bumped. The bijections  $S$  and  $S^{-1}$  can help us with this question.

To put the problem formally, we consider the random quantity defined by setting

$$N = |\{j \in \{2, 3, \dots, k\} \text{ such that } j \text{ gets bumped}\}|,$$

where we use  $|B|$  to denote the cardinality of the finite set  $B$ . Here, of course,  $N$  is a random variable that depends on the random ordering  $\tau$  of  $[n]$ , and  $N$  can take on any of the integer values  $m$  with  $0 \leq m \leq k - 1$ .

We would like to determine the probability mass function of  $N$ ; that is, we would like to find the probability of the event  $\{N = m\}$  for each  $m$ . If we reason as before, we have the logical equivalence:

$$N = m \Leftrightarrow \text{in } \tau \text{ there are exactly } m \text{ elements of } \{2, 3, \dots, k\} \text{ that precede 1.}$$

Now, in a random ordering of  $[k]$ , the value 1 occurs with the same probability at each of the  $k$  possible places, so the probability that 1 has rank  $m + 1$  is the same for all



$0 \leq m < k$ . Hence,  $N$  is uniformly distributed on the set  $\{0, 1, \dots, k-1\}$ , or, to be explicit, we have

$$P(N = m) = \frac{1}{k} \quad \text{for all } 0 \leq m < k. \quad (5)$$

Despite our long familiarity with this result, it still seems remarkable to us that the probability  $P(N = m)$  does not depend on  $m$  or  $n$  except through the minimal requirement that  $0 \leq m < k \leq n$ .

As a sidebar, we should also note that (5) implies that the random variable  $N$  has expectation

$$E[N] = \sum_{m=0}^{k-1} m \times \frac{1}{k} = \frac{k(k-1)}{2} \times \frac{1}{k} = (k-1)/2.$$

In fact, we could have guessed this from the beginning. The set of co-workers  $\{2, 3, \dots, k\}$  has cardinality  $k-1$  and each of these co-workers has probability  $1/2$  of being bumped. Linearity of expectation then tells us that the expected number of bumped co-workers has to be  $(k-1)/2$ , just as one would confirm by calculation from (5).

## Noting records: Counting cycles

There is one further topic that is too striking to be left untouched, even if the treatment we give here must be brief. Perhaps the most classical application of the operation of erasing all of the back-to-back parentheses is to determine the mean and variance—or even the distribution—of the number of cycles in a random permutation.

Remarkably enough, it will be useful to introduce the *third* recipe for constructing a cycle representation from the one-line representation of a permutation  $\sigma$ . First, we find the set of cycles as we have done before. Since we can start a cycle with any of the values in the cycle, we now choose to start each cycle with its *largest* value. Finally, we list the cycles in the order that puts leading elements into an *increasing* sequence. For example, if we take our favorite permutation  $\sigma = [3, 6, 5, 8, 1, 2, 4, 7]$  in its one-line representation, then our three recipes give us the following strings:

$$\begin{aligned} \text{First Recipe: } \sigma &= (1, 3, 5)(2, 6)(4, 8, 7), \\ \text{Second Recipe: } \sigma &= (3, 5, 1)(6, 2)(8, 7, 4), \\ \text{Third Recipe: } \sigma &= (5, 1, 3)(6, 2)(8, 7, 4). \end{aligned} \quad (6)$$

Here one should note that each of these three strings represents exactly the same permutation  $\sigma$ .

Now, to construct a mapping from  $\mathbb{S}_n$  to  $\mathbb{S}_n$ , we again erase all of the back-to-back parentheses in the third cycle representation. Thus, from the three representations of the permutation  $\sigma$  given above, we find three different images:

$$\begin{aligned} \text{Using the First Recipe: } F(\sigma) &= [1, 3, 5, 2, 6, 4, 8, 7], \\ \text{Using the Second Recipe: } S(\sigma) &= [3, 5, 1, 6, 2, 8, 7, 4], \\ \text{Using the Third Recipe: } T(\sigma) &= [5, 1, 3, 6, 2, 8, 7, 4]. \end{aligned}$$

Finally, one argues as before to show that the third mapping  $T$  is really a bijection of  $\mathbb{S}_n$  into itself. To underscore the importance of this check, one should recall that  $S$  is a bijection, but  $F$  is not!

So, how can the new mapping  $T$  help us with our declared goal of determining the mean and variance of the number of cycles in a random permutation? The key is that it gives us a way to relate the number of cycles in  $\sigma$  to another notable object, the number of *records* in the image permutation  $\tau = T(\sigma)$ .

To make this concrete, we introduce random variables  $R_i$ ,  $i = 1, 2, \dots, n$  that we define by setting  $R_i = 1$  or  $R_i = 0$  accordingly as  $a_i$  is a record maximum or not, as one scans the random permutation (written in one-line notation)  $[a_1, a_2, \dots, a_n]$  from left to right. Naturally we always have  $R_1 = 1$  since the first value  $a_1$  is always the largest one that we have seen so far. More generally, if we continue with our familiar example,  $\tau = T(\sigma) = [5, 1, 3, 6, 2, 8, 7, 4]$ , then we find that  $R_1, R_4$ , and  $R_6$  are equal to one, and the rest of the  $R_i$  are all equal to zero.

Now we come to a key observation: the number of cycles in the permutation  $\sigma$  is equal to the total number of record maxima in permutation  $\tau = T(\sigma)$ . Thus, the bijection  $T$  shows us that as random variables the number of cycles and the number of record maxima have exactly the same distribution; or, in the language of combinatorics, it shows us that the number of permutations with exactly  $k$  cycles is equal to the number of permutations with exactly  $k$  records. The punch line is that instead of always working with the number of cycles in a random permutation we can just as well work with the number of records. Moreover, from a probabilistic point of view records turn out to be almost magically nice.

First of all, for all  $1 \leq k \leq n$ , we have the notably simple relation

$$E[R_k] = P(R_k = 1) = 1/k.$$

This identity follows from the reasoning we have used before several times. Specifically, if  $R_k = 1$  then the  $k$ th item in the list has to be the largest among the first  $k$ . But the probability that a given element in a list of  $k$  elements occupies any particular place in a random ordering of that list (say the last place) is equal to  $1/k$ .

We noted earlier that the number  $C_n$  of cycles of a random permutation  $\sigma$  is equal to the number of maxima that one observes when scanning a random permutation  $\tau = T(\sigma) = [a_1, a_2, \dots, a_n]$ , so in symbols we have the identity

$$C_n = R_1 + R_2 + \dots + R_n.$$

Now, if we take the expectation of both sides of the identity, then by the linearity of expectation (cf. [2], p. 206) we find

$$E[C_n] = 1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n} \stackrel{\text{def}}{=} H_n \sim \log n \quad \text{as } n \rightarrow \infty,$$

and we have the charming appearance of the harmonic number and the logarithm in a problem that begins with counting cycles.

Moreover, one can go much further. The random variables  $\{R_i : 1 \leq i \leq n\}$  are actually independent! At first glance this assertion may seem counterintuitive, but the independence property can be proved without much work; it is a great exercise. Once one has independence, the flood gates of probability theory are open. In particular, we know that for independent random variables, the variance of the sum is equal to the sum of the variances (e.g., [2], p. 278), so we get another harmonic number formula:

$$\text{Var}[C_n] = \sum_{k=1}^n \text{Var}[R_k] = \sum_{k=2}^n \frac{1}{k} \left(1 - \frac{1}{k}\right) = H_n - \sum_{k=1}^n \frac{1}{k^2} \sim \log n \quad \text{as } n \rightarrow \infty.$$

Furthermore, by Lindeberg's classic version of the central limit theorem (cf. [2], p. 359), one then has for all real values  $t \in (-\infty, \infty)$  that

$$\lim_{n \rightarrow \infty} P\left(\frac{C_n - \log n}{\sqrt{\log n}} \leq t\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t e^{-x^2/2} dx.$$

Among other things, this formula tells us that for large  $n$ , one has probability close to 95% that the number of cycles (or the number of records) in a random permutation of  $[n]$  is within plus or minus  $2\sqrt{\log n}$  of  $\log n$ . This is surely something that one could not have easily guessed just given the raw first facts about permutations. To be sure, this excursion into the probability of random permutations has gone quickly; but, even such a brief sketch may help to give some sense of the unanticipated power that is cradled in the third recipe.

### The rest of the road: What's next?

Our *cri de cour* is simple: "Want insight? Consider a bijection!" This puts one onto a major path of modern combinatorics where one sees many ways that bijections help us to understand combinatorial structures more deeply.

To give an example of a theme worth exploration, one can consider the behavior of cycles in *special subsets of permutations*. For example, consider the set  $D_n$  of derangements of  $[n]$ , i.e., the set of permutations without fixed points. If you choose  $\sigma$  at random from  $D_n$ , what is the probability that 1 and 2 are in the same cycle? This is a problem that can be addressed with tools of the kind that we have developed here.

For a more sophisticated problem with a similar flavor, one can consider the following experiment. Choose independently two cycles of size  $n$ , say, for example,  $\sigma = (a_1, a_2, \dots, a_n)$  and  $\sigma' = (b_1, b_2, \dots, b_n)$ . Next, consider the permutation defined by the compositional product of these permutations, i.e., consider the mapping  $i \mapsto \sigma(\sigma'(i))$ . Stanley [7] found that the probability  $p_n$  that 1 and 2 are in the same cycle of this permutation is given by

$$p_n = \begin{cases} \frac{1}{2} & \text{if } n \text{ is odd} \\ \frac{1}{2} - \frac{2}{(n-1)(n+2)} & \text{if } n \text{ is even.} \end{cases}$$

Thus, for large  $n$ , we see that the cycle behavior of the permutations generated by this experiment and the cycle behavior of simple random sampling from  $\mathbb{S}_n$  are quite similar—at least with respect to this one particular question. One can anticipate that there are many other sampling experiments for  $\mathbb{S}_n$  where analogous behavior can be found.

A little further afield, one of the most famous results in combinatorial probability is Spitzer's identity [5], and it also has at its heart a bijection between  $\mathbb{S}_n$  and  $\mathbb{S}_n$ . In fact, simple modifications of the construction of Spitzer and Bohnenblust lead one to a rich family of bijections from  $\mathbb{S}_n$  and  $\mathbb{S}_n$  (cf. [8]).

Finally, we should mention the Robinson–Knuth–Schensted correspondence, which is surely the most notable bijection in the theory of permutations. The RKS correspondence is a bijection between permutations and pairs  $(P, Q)$  of remarkable discrete objects called standard Young tableaux. A clear and gentle introduction to the RKS correspondence is given by Stanton and White [4], and much further material is given in texts of Bóna [1] and Stanley [6]. Finally, we should also note that Romik [3] gives a wonderful account of the RKS correspondence and its connection to the longest increasing subsequence problem, which is rich topic that has experienced stunning progress over recent years.

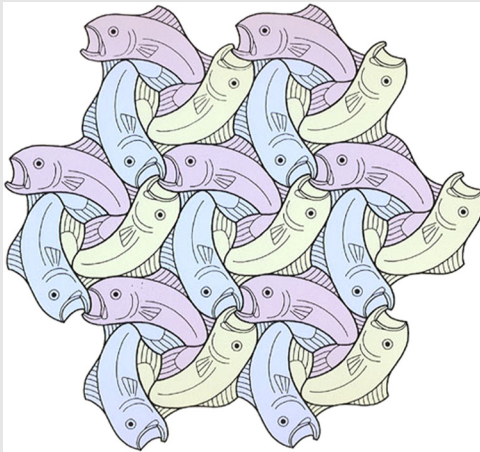
## REFERENCES

1. M. Bóna, *Combinatorics of Permutations*. Chapman and Hall/CRC, New York, 2004.
2. P. Billingsley, *Probability and Measure*. John Wiley and Sons, New York, 1995.
3. D. Romik, *The Surprising Mathematics of Longest Increasing Subsequences*. Cambridge Univ. Press, Cambridge, 2015.
4. D. Stanton, D. White, *Constructive Combinatorics*. Springer-Verlag, New York, 1986.
5. F. Spitzer, A combinatorial lemma and applications to probability, *Trans. Amer. Math. Soc.* **82** (1956) 323–339.
6. R. P. Stanley, *Enumerative Combinatorics, Vol. I*. Cambridge Univ. Press, New York, 1997.
7. ———, Two enumerative results on cycles of permutations, *European J. Combin.* **32** (2011) 937–947.
8. J. M. Steele, The Bohnenblust–Spitzer algorithm and its applications, *J. Comput. Appl. Math.* **142** (2002) 235–249.
9. I. Tomescu, *Problems in Combinatorics and Graph Theory*. Wiley-Interscience, New York, 1985.

**Summary.** The random—or orderly—seating of passengers on a bus is used to motivate several questions about cycles of permutations. These in turn motivates the investigation of bijections between special subsets of permutations. The goal, of course, is to give simple explanations of surprising facts.

**VLADIMIR POZDNYAKOV** (MR Author ID: [650411](#)) received his Ph.D. in statistics in 2001 from the University of Pennsylvania. Since that time he has taught at the University of Connecticut, where he is currently professor of statistics. His research is mostly in applied probability, and he has a particular interest in the discovery and exploitation of martingale tricks.

**J. MICHAEL STEELE** (MR Author ID: [166615](#)) received his Ph.D. in mathematics from Stanford University. He has taught at U.B.C., Stanford, CMU, Princeton, and the Wharton School of the University of Pennsylvania. He enjoys mathematical inequalities and sequential decision making, and he pays the rent with stochastic calculus and financial time series.



### Artist Spotlight Robert Fathauer

*Three Fishes*, Robert Fathauer; limited edition screen print, 1994. This work depicts tessellating fish having threefold rotational symmetry about the tail, top fin, and mouth. The motif bears a resemblance to one of Escher's tessellation designs, but the symmetry of the design is quite different. Escher's design has fourfold rotational symmetry about the tail and the top fin, with twofold rotational symmetry about the jaw.

See interview on page 220.