# Notes on Permutations: Cycles and Transpositions

Tyler Neylon

332.2022

[Formats: html | pdf]

This is a collection of notes about permutations — all the different ways of ordering a set of distinct elements. In group theory, the group of permutations of size $n$ is denoted $S_n$. For example:

| | | | | | | |
|---|---|---|---|---|---|---|
| $S_2$ | 12 | 21 | | | | |
| $S_3$ | 123 | 132 | 213 | 231 | 312 | 321 |
| $S_4$ | 1234 | 1243 | 1324 | 1342 | 1423 | 1432 |
| | 2134 | 2143 | 2314 | 2341 | 2413 | 2431 |
| | 3124 | 3142 | 3214 | 3241 | 3412 | 3421 |
| | 4123 | 4132 | 4213 | 4231 | 4312 | 4321 |

You can think of a single permutation as either a particular ordering of the integers 1 through $n$, or, equivalently, as a 1-1 and onto mapping $f : \{1, \ldots, n\} \to \{1, \ldots, n\}$. There's a natural way to write such a mapping $f$, which is to list out $f(1), f(2), \ldots, f(n)$ as a *string*; you can consider the lists of elements of $S_2, S_3, S_4$ above to be in this string notation. By default, this article will view permutations as mappings.

This article covers the following independent observations about permutations:

- I'll introduce the idea of a *cut-merge* operation, and show how this corresponds to a *transposition*; a transposition is a permutation that exchanges two elements of a set, and changes no others. Given the set {1, 2, 3, 4, 5}, the swap of 2 and 5 would be a transposition; I might write this permutation as 15342.
- Permutations can be split into two kinds: *odd* permutations and *even* permutations. Just as odd + even = odd when it comes to integer addition, the analogous rules hold for permutations under composition. For example, if I apply an odd permutation, and then apply an even permutation after that, the result of the combined permutations is itself an odd permutation (and this analogy holds for all combinations of odd/even). I'll present a simple proof of this breakdown using cut-merge operations.

- I'll discuss a generalization of the parity (even-ness or odd-ness) of a permutation $\pi$ that I call the *magnitude* $m(\pi)$ of the permutation.
- I'll define the *cycle structure* $\mathsf{cs}(\pi)$, which captures the sizes of cycles in the permutation $\pi$; and I'll show that $\mathsf{cs}(a \cdot b) = \mathsf{cs}(b \cdot a)$ for any two permutations $a, b$.
- I'll show that the magnitude $m(\cdot)$ acts like a norm on the group $S_n$, and that it can define a coherent distance function $\mathsf{dist}(\cdot, \cdot)$ on $S_n$.

When I say these are *independent* observations, I mean that these ideas are new to me. I have not tried much to check if they are new to the world altogether — in all likelihood, many of these ideas have been discovered by others before me. Still, I think they're fun ideas worth sharing.

# 1 Cycle Notation

Any permutation is a specific ordering of a set of distinct elements. This article focuses on *finite* permutations, and I'll generally use the variable $n$ to denote the size of the set being permuted. It's simple and traditional to think of a permutation as a bijection (a 1-1 and onto mapping) of the first $n$ integers — for example, here's a permutation of the set $\{1, 2, \ldots, 8\}$ in string notation:

$$\pi = 15273648. \tag{1}$$

Despite the possible confusion with the constant $\pi = 3.141\ldots$, it's also traditional to use the variable $\pi$ for a permutation, as (clearly) $\pi$ stands for "$\pi$ermutation."

There are many fun ways to denote or think about a particular permutation. For example, given $\pi$ as in (1), we could think of it as the sequence $\pi_1 = 1$, $\pi_2 = 5$, $\ldots$, $\pi_8 = 8$. We could also think of it as a function $\pi : [n] \to [n]$, using $[n]$ to denote the integers $\{1, 2, \ldots, n\}$.

This article uses a particular notation for permutations called *cycle notation*: the idea for cycle notation is to write out a string of integers in $[n]$ with the interpretation that any consecutive pair $ij$ indicates that $i$ is mapped to $j$, and any $i$ at the end of a parenthesized group maps to $j$ at the start of the same group. Our permutation from (1) looks like this in cycle notation:

$$\pi = (2\ 5\ 3)(4\ 7)$$

because $\pi(2) = 5$, $\pi(5) = 3$, and $\pi(4) = 7$. Each sequence in parentheses *wraps around*, so $\pi(3) = 2$ and $\pi(7) = 4$. Thus each parenthesized sequence is a *cycle*. It is understood that any omitted elements are mapped to themselves; since $\pi(1) = 1$, we don't need to include 1 in the cycle notation for $\pi$.

Examples:

| | | | |
|---|---|---|---|
| For $n = 5$, | (1 3) | denotes | `32145`. |
| For $n = 5$, | (2 3)(4 5) | denotes | `13254`. |
| For $n = 8$, | (1 5 2)(3 7)(4 8 6) | denotes | `51782436`. |

You could write the same permutation many ways in cycle notation; for example:

$$\pi = (7\ 4)(5\ 3\ 2) = (3\ 2\ 5)(7\ 4) = (2\ 5\ 3)(4\ 7).$$

Among those choices, I consider the last one to be standard because it's lexicographically first.

## 2 The Cut-Merge Operation

Now I'll introduce an operation that acts naturally on the cycle notation of a permutation. Intuitively, a *merge* is an operation that combines two cycles by concatenating their cycle notation. When two elements — say 3 and 5 — are in different cycles, then I'll write $\pi * (3\ 5)$ to mean "merge, in $\pi$, the cycles starting with 3 and 5."

An example:

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

What if 3 and 5 were already in the same cycle? Then $\pi * (3\ 5)$ takes on the meaning "*cut* apart the subcycles starting with 3 and 5:"

$$(1\ 2)(3\ 4\ 5\ 6) * (3\ 5) = (1\ 2)(3\ 4)(5\ 6)$$

or

$$(1\ 2\ 3\ 4\ 5) * (3\ 5) = (3\ 4)(5\ 1\ 2);$$

this last example might be easier to see when you keep in mind that

$$(1\ 2\ 3\ 4\ 5) = (\underline{3}\ 4\ \underline{5}\ 1\ 2).$$

Now I'm ready to more carefully define the cut-merge operation $*(x\ y)$. In the definition below, I'll write $(x\cdots)$ to denote the cycle starting with $x$; and, analogously, I'll write $(x\cdots\ y\cdots)$ to indicate a cycle containing both $x$ and $y$, possibly with other elements indicated by the dots.

**Definition** The *cut-merge* operation $*(x\ y)$ operates on a permutation $\pi$ via

$$\pi * (x\ y) = \begin{cases} (x\cdots\ y\cdots) & \text{if } \pi = (x\cdots)(y\cdots) \text{ or,} \\ (x\cdots)(y\cdots) & \text{if } \pi = (x\cdots\ y\cdots). \end{cases}$$

If there are any other cycles (ones without $x$ or $y$), then they are unaffected by the $(x\ y)$ operation.

The operation *cuts* a cycle if $x$ and $y$ are in the same cycle; otherwise it *merges* the cycles of $x$ and $y$. A shorthand for the definition is:

$$(x\cdots\ y\cdots) \underset{*(x\ y)}{\longleftrightarrow} (x\cdots)(y\cdots).$$

Note that a merge can still operate on elements not explicitly written in cycle notation. If I write $e$ to denote the identity permutation, with $\pi_i = i\ \forall i$, then

$$e * (x\ y) = (x\ y).$$

Similarly,

$$(1\ 2) * (2\ 3) = (2\ 1\ 3)$$

because, setting $x = 2$ and $y = 3$, we think of the left side as $(x\cdots)(y\cdots) = (21)(3)$.

**Observation 1** The cut-merge operation $\pi * (x\ y)$ is simply $\pi$ composed with the permutation $(x\ y)$.

To state this observation again, I'll introduce the notation $\rho_{ab}()$ for the function that swaps input elements $a$ and $b$, but otherwise acts as the identity function. Thus $\rho_{ab}()$ is the function version of the *transposition* that we would write in cycle notation as $(a\ b)$. With $\rho_{ab}()$ defined, I can express observation 1 like so:

$$(\pi * (x\ y))\,(i) = \rho_{xy}(\pi(i)), \tag{2}$$

where I'm thinking of permutations as functions, so the $(i)$ sub-expression is evaluating the function on input $i \in \{1, 2, \ldots, n\}$.

I could stop to carefully prove observation 1, but I think you'll have more fun if you convince yourself it's true on your own.

Why did I bother to define the cut-merge operation when I could have just started with (2) instead? Because the perspective of the definition I'm using will let us make some interesting further observations, as we'll see below.

# 3 The Magnitude and Parity of a Permutation

**Definition** The *magnitude* of a permutation $\pi$ is given by

$$m(\pi) := \#(\text{cycle elements of } \pi) - \#(\text{cycles of } \pi),$$

where a *cycle element* of $\pi$ is any element written in the cycle notation of $\pi$, and where we similarly count the *cycles* of $\pi$ based on how many cycles are written in $\pi$'s cycle notation.

For example, the permutation $\pi = (2\ 5\ 3)(4\ 7)$ has 5 cycle elements and 2 cycles, so $m(\pi) = 3$.

Notice that $m(\pi)$ remains the same if we include singleton cycles in our cycle notation. It would be nonstandard to include singletons, but it would still be a consistent notation. Continuing our example, $\pi = (1)(2\ 5\ 3)(4\ 7)(6) \Rightarrow m(\pi) = 7 - 4 = 3$, as before. This leads to:

**Observation 2**

$$m(\pi) = n - \#(\text{orbits}),$$

where an *orbit* is either a cycle with multiple elements, or a *singleton*, which is an element $i : \pi(i) = i$.

Let's see why the definition of $m(\pi)$ is interesting.

**Observation 3**

(i). Every merge increases $m(\pi)$ by 1.
(ii). Every cut decreases $m(\pi)$ by 1.
(iii). $m(\pi)$ is the least number of cut-merge operations which can reach $\pi$ from the identity permutation $e$.

Parts (i) and (ii) are easy to check.

Part (iii) is clear from the cut-merge perspective in that (a) we cannot build $\pi$ in fewer operations, based on (i); and (b) we can indeed build $\pi$ in $m(\pi)$ merges by appending together the appropriate singletons one at a time. For example:

$$(2\ 5\ 3)(4\ 7) = e * (2\ 5) * (2\ 3) * (4\ 7). \tag{3}$$

Starting now, I'll write simply $\sigma\tau$ or $\sigma \cdot \tau$ to denote the composition of permutations $\sigma$ and $\tau$. Using observation 1, we can see that the construction pattern used in (3) must work for any cycle that we want to write as a composition of transpositions:

$$(x_1\ x_2 \cdots x_k) = (x_1\ x_2)(x_1\ x_3) \cdots (x_1\ x_k).$$

In other words, every permutation is the product of $m(\pi)$ transpositions, and cannot be the product of fewer. It could be the product of more, though:

$$(1\ 2\ 3) = (1\ 2)(1\ 3)(8\ 9)(8\ 9).$$

Now we're ready to define an interesting way to classify all permutations into two different types:

| **Definition** | A permutation $\pi$ is *odd* if $m(\pi)$ is odd, and *even* otherwise. |
|---|---|

As soon as we call something even or odd, we expect some kind of behavior like this to happen:

$$\begin{aligned} \text{even} + \text{even} &= \text{even} \\ \text{odd} + \text{even} &= \text{odd} \\ \text{odd} + \text{odd} &= \text{even}. \end{aligned} \tag{4}$$

It turns out that these rules do indeed hold true in the sense that, for example, if we compose an odd permutation with another odd permutation, the result is an even permutation. Next I'll explain why the intuitive identities of (4) are true for permutations.

The key to the proof is to see that, for any sequence of transpositions $t_1, t_2, \ldots, t_k$:

$$m(t_1 t_2 \cdots t_k) \text{ is even iff } k \text{ is even.} \tag{5}$$

This fact follows from observation 3. In more detail: Certainly $m(e) = 0$ (recall that $e$ is the identity permutation); this corresponds to $k = 0$ in (5), an empty product. Now for larger values of $k$: If $\pi = t_1 t_2 \cdots t_{k-1}$ and we compose $\pi$ with $t_k$ (move from $\pi$ to $\pi \cdot t_k$), then the parity (even/odd-ness) of the permutation switches because $m(\pi \cdot t_k) = m(\pi) \pm 1$ by observation 3. So $\pi \cdot t_k$ is even iff $\pi$ is odd. This is enough to complete a proof of (5) by induction on $k$.

It will be useful to define the sign function of a permutation $\pi$ like so:

$$\mathsf{sign}(\pi) := \begin{cases} 1 & \text{if } \pi \text{ is even, and} \\ -1 & \text{if } \pi \text{ is odd.} \end{cases}$$

Notice that $\mathsf{sign}(\pi) = (-1)^{m(\pi)}$.

When we use this sign function, we're essentially moving from an additive view of even/odd-ness and into a multiplicative view — but the two views are equivalent. For example, the equation

$$\text{even} + \text{odd} = \text{odd}$$

becomes

$$(-1)^{\text{even}}(-1)^{\text{odd}} = (-1)^{\text{odd}}.$$

The general equation $(-1)^a(-1)^b = (-1)^{a+b}$ provides the equivalence.

We can now state the idea of (4) both more formally and more succinctly.

For any two permutations $\sigma$ and $\tau$,

$$\mathsf{sign}(\sigma \cdot \tau) = \mathsf{sign}(\sigma) \cdot \mathsf{sign}(\tau). \tag{6}$$

Let's see why (6) is true. Let $k = m(\sigma)$ and $k' = m(\tau)$. Then, using observation 3(iii), we can find transpositions $(t_i)_1^k$ and $(t'_i)_1^{k'}$ so that

$$\begin{aligned}\sigma &= t_1 \cdots t_k \\ \tau &= t'_1 \cdots t'_{k'}.\end{aligned}$$

Then

$$\begin{aligned}\mathsf{sign}(\sigma \cdot \tau) &= (-1)^{m(\sigma\tau)} && \text{by def'n of } \mathsf{sign} \\ &= (-1)^{m(t_1 \cdots t_k \cdot t'_1 \cdots t'_{k'})} && \text{by def'n of } t_i \text{ and } t'_i \\ &= (-1)^{k+k'} && \text{by (5)} \\ &= (-1)^k (-1)^{k'} \\ &= \mathsf{sign}(\sigma)\mathsf{sign}(\tau).\end{aligned}$$

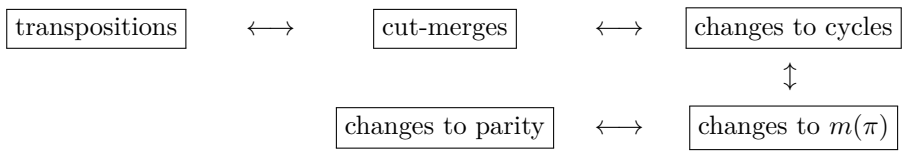
This completes the proof of (6). And (6) expresses the same idea as (4), so that we have now justified all of those formulae.

Moving back to group theory: Notice that if $\sigma$ and $\tau$ are both even permutations, then so is $\sigma\tau$. In other words, the subset of even permutations is closed under composition, so that they form a subgroup of $S_n$. This subgroup is called the *alternating group*. Since odd permutations are not closed under composition, they don't form a subgroup.

# 4 Previous Approaches to Permutation Parity

I like the cut-merge proof of (6) because it provides an intuition for *why* parity makes sense for permutations. In a nutshell, every permutation is a sequence of transpositions (cut-merges), and every transposition is a flip of the parity because the magnitude must either increase or decrease by 1.

Here's a sketch of the intuitive paths we've crossed:




I'll contrast the proof above of (6) with two other approaches.

## 4.1 Herstein's Proof

In his book *Topics in Algebra*, I.N. Herstein uses the Vandermonde polynomial:

$$p(x_1, \ldots, x_n) := \prod_{i<j} (x_i - x_j).$$

Given a permutation $\pi$, we can now define

$$\pi(p(x_1, \ldots, x_n)) := p(x_{\pi_1}, \ldots, x_{\pi_n}).$$

The polynomial $\pi(p)$ is still a product of $(x_i - x_j)$ for all $i \neq j$, but the sign of the polynomial may change. For example, when $n = 3$, and $\pi = (1\ 2)$:

$$\begin{aligned} p(x_1, x_2, x_3) &= (x_1 - x_2)(x_1 - x_3)(x_2 - x_3) \\ \pi(p) &= (x_2 - x_1)(x_2 - x_3)(x_1 - x_3) \\ &= -p. \end{aligned}$$

Since $\pi(p) = \pm p$ for any $\pi$, we can (re)define

$$\mathsf{sign}(\pi) := \pi(p)/p \in \{-1, +1\}. \tag{7}$$

(Temporarily forget our original definition of $\mathsf{sign}(\pi)$. We'll re-prove (6), and because our new and old $\mathsf{sign}()$ functions are the same on $e$ and on transpositions, it will follow that they are in fact the same function.)

Notice that

$$\pi(-p) = -\pi(p) \tag{8}$$

based on the definition of $\pi(p)$.

This means that, for any two permutations $\sigma$ and $\tau$,

$$\begin{aligned} \mathsf{sign}(\sigma\tau) \cdot p &= \tau(\sigma(p)) && \text{by (7)} \\ &= \tau(\mathsf{sign}(\sigma) \cdot p) && \text{by (7)} \\ &= \mathsf{sign}(\sigma) \cdot \tau(p) && \text{by (8)} \\ &= \mathsf{sign}(\sigma) \cdot \mathsf{sign}(\tau) \cdot p. \end{aligned}$$

In summary, $\mathsf{sign}(\sigma \cdot \tau) = \mathsf{sign}(\sigma) \cdot \sigma(\tau)$, which confirms (6) for this new definition of $\mathsf{sign}()$. This proof provides less insight than the cut-merge approach as to *why* (6) is true — into what is going on behind the scenes to make things click.

## 4.2 Artin's Proof

The other approach I'll consider is from the book *Algebra* by Michael Artin. This second alternative proof works by translating permutations into the language of linear algebra, and relies upon properties of the determinant of a matrix. On the off chance that you're not fluent in properties of determinants, you can skim or skip this section and safely resume reading afterwards (the remainder of this note does not rely on linear algebra). I'm not a huge fan of this proof because it's so easy to get lost in the details of tracking the relationship between vector coordinates and how permutations act on them. I'll break the proof down into small steps to walk you through the dance of indices.

I'll use the variable $e_i$ to denote the column vector of length $n$:

$$e_i = \begin{matrix} 0 \\ 0 \\ \vdots \\ 1 \longleftarrow i^{\text{th}} \text{ place.} \\ \vdots \\ 0 \\ 0 \end{matrix}$$

For example, if $n = 3$, then

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Now I can define, for permutation $\pi$, the $n \times n$ matrix $M_\pi$ as having $e_{\pi_i}$ for its $i^{\text{th}}$ column; this is how Artin represents a permutation as a matrix. For example, if $\sigma = (3\ 2\ 1)$, then $M_\sigma = \begin{pmatrix} & & 1 \\ & 1 & \\ 1 & & \end{pmatrix}$.

Let's see why this matrix definition makes sense. Using the column-expansion perspective of matrix multiplication, with $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$,

$$M_\pi x = x_1 \cdot e_{\pi_1} + x_2 \cdot e_{\pi_2} + \ldots + x_n \cdot e_{\pi_n} = \begin{pmatrix} \vdots \\ x_i \\ \vdots \end{pmatrix} \longleftarrow \text{ in row } \pi_i. \quad (9)$$

In other words, when we multiply $M_\pi$ on the left of a column vector, we take whatever value was in the $i^{\text{th}}$ row and move it to the $\pi_i^{\text{th}}$ row. For example, $M_\sigma \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} & & 1 \\ & 1 & \\ 1 & & \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix}$. Continuing with this intuition, left-multiplying

by $M_\tau M_\sigma$ is like performing $\sigma$ followed by $\tau$ in the sense that the $i^{\text{th}}$ row is moved to the $\tau(\sigma(i))^{\text{th}}$. I personally prefer to have statements like this algebraically checked, so in the next couple of paragraphs, I'll spell out more carefully why this is true.

It's useful to slightly rewrite (9) as in the right-hand expression below, using the notation $[v]_i$ to indicate the $i^{\text{th}}$ coordinate of column vector $v$:

$$\left[ M_\pi \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \right]_{\pi_i} = x_i \quad \Rightarrow \quad [M_\pi x]_i = x_{\pi^{-1}i}. \tag{10}$$

So left-multiplying by both $M_\tau$ and $M_\sigma$ works like so, using (10):

$$[M_\tau M_\sigma x]_i = [M_\sigma x]_{\tau^{-1}(i)} = x_{\sigma^{-1}(\tau^{-1}(i))}. \tag{11}$$

Writing $M_{\sigma\tau}$ for the matrix of permutation $\sigma\tau$, we have

$$[M_{\sigma\tau} x]_i = x_{(\sigma\tau)^{-1}(i)} = x_{\sigma^{-1}(\tau^{-1}(i))}. \tag{12}$$

Together, (11) and (12) confirm that:

$$M_{\sigma\tau} = M_\tau M_\sigma$$

in the sense that they act the same in matrix multiplications, from which we can conclude that they are in fact identical matrices.

At long last, I'm ready to (re)define the $\mathsf{sign}()$ function, following Artin:

$$\mathsf{sign}(\pi) := \det(M_\pi),$$

where det() is the determinant function of a square matrix.

It turns out that this definition is also consistent with our earlier definitions because:

- $\det(M_t) = -1$ for any transposition $t$ because exchanging any two columns of a matrix negates its determinant; and
- when $\pi = \prod_{i=1}^{k} t_i$ for transpositions $(t_i)_1^k$, $\mathsf{sign}(\pi) = \det(M_\pi) = \det(\prod_{i=1}^{k} M_{t_i}) = (-1)^k$ = our cut-merge definition of $\mathsf{sign}()$, based on (5).

In using this approach, Artin is able to show that (6) holds simply by delegating the work to the properties of matrix determinants. In particular, using the fact that $\det(AB) = \det(A)\det(B)$:

$$\mathsf{sign}(\sigma \cdot \tau) = \det(M_\tau M_\sigma) = \det(M_\tau)\det(M_\sigma) = \mathsf{sign}(\sigma)\mathsf{sign}(\tau).$$

Artin's proof is certainly valid, though I think it offers less insight into *why* equation (6) holds when compared to the cut-merge proof.

# 5 Cycle Structures

Let $\sigma = (1\ 3)(2\ 4\ 5)$ and $\tau = (1\ 4)(2\ 3)$. Then

$$\sigma\tau = (1\ 2)(3\ 4\ 5), \text{ and}$$
$$\tau\sigma = (1\ 5\ 2)(3\ 4).$$

Notice that these two permutations each have one cycle of length 2 and another of length 3. Is this a coincidence? It turns out that this similarity of structure will *always* hold between $\sigma\tau$ and $\tau\sigma$. I'll have to define a new concept before I can precisely formulate this relationship.

**Definition** Given a permtuation $\pi$, a *cycle set* $S$ of $\pi$ is a set of elements in the domain of $\pi$ such that, for any $a, b \in S$, $\exists k : \pi^k(a) = b$. Notice that every element in the domain of $\pi$ is in exactly one cycle set of $\pi$.

The *cycle structure* $\mathsf{cs}(\pi)$ of $\pi$ is a function $f : \mathbb{N}_{\geq 1} \to \mathbb{N}_{\geq 0}$ such that $f(i) :=$ the number of cycle sets of $\pi$ with $i$ elements.

As a shorthand, I'll write $(f(1), f(2), f(3), \ldots)$ to express a cycle structure. For example, when $n = 5$, the permutation $(1\ 2)(3\ 4\ 5)$ has cycle structure $(0, 1, 1)$; so does the permutation $(1\ 5\ 2)(3\ 4)$. Now I can state:

**Observation 4** For any permutations $\sigma$ and $\tau$,

$$\mathsf{cs}(\sigma\tau) = \mathsf{cs}(\tau\sigma).$$

I'll provide two proofs because they each shed their own light on why this is true.

**Proof 1** Given permutation $\pi$, I'll write $\mathsf{order}_\pi(i) = k$ to mean that $\pi^k(i) = i$, and there is no $\ell < k : \pi^\ell(i) = i$. To prove the observation, it's sufficient to show that there is a 1-1 and onto mapping $f : [n] \to [n]$ such that $\mathsf{order}_{\sigma\tau}(i) = k \Leftrightarrow \mathsf{order}_{\tau\sigma}(f(i)) = k$.

Suppose $\mathsf{order}_{\sigma\tau}(i) = k$. Let $j = \tau^{-1}(i)$. In the following equations, I'll use a left-to-write, input-to-output notation, where $i\pi$ means $\pi(i)$ and $i\sigma\tau$ means $\tau(\sigma(i))$. We have:

$$i = i(\sigma\tau)^k = j\tau(\sigma\tau)^k = j(\tau\sigma)^k\tau.$$

Now apply $\tau^{-1}$ to the left-most and right-most expressions above to arrive at:

$$j = i\tau^{-1} = j(\tau\sigma)^k \ \Rightarrow\ \mathsf{order}_{\tau\sigma}(j) \leq k = \mathsf{order}_{\sigma\tau}(i). \tag{13}$$

On the other hand, if $\ell = \mathsf{order}_{\tau\sigma}(j)$, then

$$i = (j)\tau = (j)(\tau\sigma)^\ell\tau = (j)\tau(\sigma\tau)^\ell = i(\sigma\tau)^\ell$$
$$\Rightarrow \mathsf{order}_{\sigma\tau}(i) \leq \ell = \mathsf{order}_{\tau\sigma}(j). \tag{14}$$

Combining (13) and (14),

$$\mathsf{order}_{\sigma\tau}(i) = \mathsf{order}_{\tau\sigma}(j = \tau^{-1}(i)),$$

and it turns out that $\tau^{-1}$ is the mapping $f : [n] \to [n]$ we needed to complete the proof. □

**Proof 2** Given any permutation product $\sigma\tau$, we can decompose $\tau$ into transpositions as in $\tau = t_1 \cdots t_k$. Now we can think in terms of a series of smaller steps that carry us from $\sigma\tau$ over to $\tau\sigma$:

$$\begin{aligned}
\pi_0 &= \sigma\tau = \sigma \cdot t_1 \cdots t_k \\
\pi_1 &= t_k \cdot \sigma \cdot \tau \cdot t_k^{-1} = t_k\sigma t_1 \cdots t_{k-1} \\
&\vdots \\
\pi_k &= t_1 \cdots t_k\sigma\tau t_k^{-1} \cdots t_1^{-1} = \tau\sigma.
\end{aligned}$$

Each small step here can be characterized as $\pi_{i+1} = t_{k-i}\pi_i t_{k-i}$, using the fact that $t_j^{-1} = t_j$, which is true for any transposition. So the problem is reduced to showing that, for any permutation $\pi$ and transposition $t$,

$$\mathsf{cs}(\pi) = \mathsf{cs}(t\pi t). \tag{15}$$

It turns out that applying $t = (x\ y)$ in this manner simply swaps the places of $x$ and $y$ in the cycle notation of $\pi$. For example:

$$(1\ 3)(\underline{1}\ 2\ 5\ \underline{3}\ 7)(1\ 3) = (\underline{3}\ 2\ 5\ \underline{1}\ 7).$$

This confirms (15), which completes the proof. □

Notice that observation 4 can be generalized for any rotation of any finite product of permutations. For example,

$$\mathsf{cs}(abcd) = \mathsf{cs}(bcda) = \mathsf{cs}(cdab) = \mathsf{cs}(dabc).$$

The equation $\mathsf{cs}(\sigma\tau) = \mathsf{cs}(\tau\sigma)$ makes it tempting to suspect that cycle structure is preserved no matter what order we multiply a given list of permutations. For example, we could ask if the following identities would hold for any three permutations $a, b, c$:

$$\mathsf{cs}(abc) \stackrel{?}{=} \mathsf{cs}(acb) \stackrel{?}{=} \mathsf{cs}(bac) \stackrel{?}{=} \mathsf{cs}(bca), \text{etc.}$$

for all the orderings of $a, b, c$; and similarly for any finite list of permutations.

But this is not true in general. For example, if $a = (1\ 2\ 3)$, $b = (1\ 2)$, $c = (1\ 3)$, then

$$a\ b\ c = (1\ 2\ 3)(1\ 2)(1\ 3) = (1\ 3\ 2)$$
$$c\ b\ a = (1\ 3)(1\ 2)(1\ 2\ 3) = e.$$

There is a natural follow-up question to observation 4 which I can state clearly once I've provided a couple new definitions.

Given any permutation $\pi$, let

$$\mathsf{flips}(\pi) := \{b \cdot a : \pi = a \cdot b\}.$$

And let

$$\mathsf{same\_cycles}(\pi) := \{\sigma \in S_n : \mathsf{cs}(\sigma) = \mathsf{cs}(\pi)\}.$$

**Question 1** Observation 4 tells us that, for any $\pi \in S_n$,

$$\mathsf{flips}(\pi) \subset \mathsf{same\_cycles}(\pi).$$

Are these sets actually equal?

This is similar to asking: Given $\pi$ and the freedom to decompose it as $\pi = a \cdot b$, what are all the possible values of $b \cdot a$?

The answer arises from a simple permutation decomposition. Let's start with:

**Observation 5** The cycle notation of $\tau^{-1}\sigma\tau$ is the same as the cycle notation of $\sigma$ after each element $i$ is replaced with $(i)\tau$.

The observation is true because

$$(i\tau)(\tau^{-1}\sigma\tau) = i\sigma\tau;$$

that is, if $\sigma' := \tau^{-1}\sigma\tau$ and $\sigma : i \mapsto j$, then $\sigma' : i\tau \mapsto j\tau$.

For example, if $\tau = (2\ 5\ 3)$ and $\sigma = (1\ 2)(3\ 4\ 5)$, then

$$\tau^{-1}\sigma\tau = (1\ 5)(2\ 4\ 3) = [(1\ 2)(3\ 4\ 5)].\text{replace}(\tau).$$

Note that this replacement operation is specific to cycle notation. If we were to instead write out $\pi$ as the string $\pi_1\pi_2 \ldots \pi_n$, then the corresponding string for $\tau^{-1}\sigma\tau$ is no longer replacement with $\tau$. I'll illustrate this distinction with an example:

$$\begin{aligned}
\tau &= (2\ 5\ 3) \\
\sigma &= 2\ 1\ 4\ 5\ 3 \quad (\text{This is } \sigma_1\sigma_2\sigma_3\sigma_4\sigma_5.) \\
\tau^{-1}\sigma\tau &= 5\ 4\ 2\ 3\ 1 \neq [2\ 1\ 4\ 5\ 3].\text{replace}(\tau) \\
\sigma\tau &= 5\ 1\ 4\ 3\ 2 = [2\ 1\ 4\ 5\ 3].\text{replace}(\tau).
\end{aligned}$$

To help build intuition for the set $\mathsf{same\_cycles}(\pi)$, we can choose a canonical "first" element of $\mathsf{same\_cycles}(\pi)$. A simple way to do this it to treat all singleton-free cycle notations in $\mathsf{same\_cycles}(\pi)$ as strings, treating the ")" character as the last in the alphabet (we'll ignore the "(" character for the sake of ordering). Then we can define $\mathsf{canon}(\pi)$ to be the lexicographically first element of $\mathsf{same\_cycles}(\pi)$. For example, $\mathsf{canon}(3\ 4\ 2)(1\ 5) = (1\ 2\ 3)(4\ 5)$. If we had not specified the ordering of ")", then it would be unclear if $(1\ 2)(3\ 4\ 5)$ or $(1\ 2\ 3)(4\ 5)$ were first.

Now we're ready for:

**Answer to Question 1** Given any $\sigma \in \mathsf{same\_cycles}(\pi)$, we can decompose $\pi = a \cdot b$ so that $b \cdot a = \sigma$. That is, $\mathsf{flips}(\pi) = \mathsf{same\_cycles}(\pi)$.

**Proof** We can use observation 5 to find $\tau$ so that

$$\pi = \tau^{-1}\sigma\tau. \tag{16}$$

Let $a = \tau^{-1}$ and $b = \tau\pi$. Then

$$\begin{aligned} ab &= \pi \\ ba &= \tau\pi\tau^{-1} = \sigma \quad \text{by (16).} \end{aligned}$$

□

The main idea of this proof is to use observation 5 to see that $\mathsf{cs}(\tau^{-1}\sigma\tau) = \mathsf{cs}(\sigma)$ for any $\tau$. As the proof shows, by choosing the right $\tau$, we can make $\tau^{-1}\sigma\tau$ equal to anything in $\mathsf{flips}(\sigma)$. And anything of the form $\tau^{-1}\sigma\tau$ is in $\mathsf{flips}(\sigma)$ by considering $a = \tau$ and $b = \tau^{-1}\sigma$.

Note that the $\tau$ used in the proof may not be unique because there are different cycle notation strings for the same permutation. Consider $\pi = (1\ 2)$ and $\sigma = (3\ 4)$. We can convert $\pi$ into $\sigma$ in different ways:

$$\begin{array}{lll} (1\ \ 2) & & (1\ \ 2) \\ \downarrow\ \downarrow & \tau = (1\ 3)(2\ 4) & \downarrow\ \downarrow \quad \tau' = (1\ 4)(2\ 3) \\ (3\ \ 4) & & (4\ \ 3) \end{array}$$

The above diagram is an informal way of writing that

$$\tau^{-1}\pi\tau = (3\ 4) = (4\ 3) = \tau'^{-1}\pi\tau'.$$

# 6 The Magnitude is Like a Norm

In this section I'll point out some interesting properties of the magnitude function.

**Observation 6** For any permutations $\pi$ and $\sigma$:

1. $m(\pi) \geq 0$; and $m(\pi) = 0$ iff $\pi = e$.
2. $m(\pi\sigma) \leq m(\pi) + m(\sigma)$.
3. $m(\pi^{-1}) = m(\pi)$.
4. $m(\pi\sigma) = m(\sigma\pi)$.

**Proof** Start by noticing that $m(\pi)$ is the least number of transpositions by which we can get from $e$ to $\pi$; ie, $m(\pi)$ is the least integer for which we can write $\pi = \prod_{i=1}^{m(\pi)} t_i$ for some transpositions $(t_i)_1^{m(\pi)}$.

Considering each of $\pi$ and $\sigma$ as a composition of this minimal number of transpositions, parts (i) and (ii) are straightforward.

Part (iii) follows since the cycle notation for $\pi^{-1}$ is simply the reverse of the cycle notation for $\pi$.

Part (iv) follows from observation 4. $\square$

In many ways, the magnitude of a permutation acts like a vector norm — the key difference is that we don't have a meaningful interpretation of $k\pi$ for any $k \in \mathbb{R}$, so we don't have the additional rule $m(k\pi) = |k| \cdot m(\pi)$.

This makes it interesting to define a notion of distance between permutations:

**Definition** Given permutations $\pi$ and $\sigma$,

$$\mathsf{dist}(\pi, \sigma) := m(\pi^{-1}\sigma).$$

I'll justify why this definition makes sense. Let $x = \pi^{-1}\sigma$. Then $\pi x = \sigma$, so $\mathsf{dist}(\pi, \sigma)$ is measuring the fewest "hops" — transpositions — needed to get from $\pi$ to $\sigma$.

We'd like a good distance function to be symmetric, and luckily for us, this one is:

$$\mathsf{dist}(\sigma, \pi) = m(\sigma^{-1}\pi) = m(\pi^{-1}\sigma) = \mathsf{dist}(\pi, \sigma),$$

the middle equality following since $m(\pi) = m(\pi^{-1})$ by observation 6. This symmetry addresses the apparent inelegance of the definition; indeed, the definition may appear to treat $\pi$ and $\sigma$ differently, but the end result does not depend on the order of these inputs to $\mathsf{dist}()$.

To conclude this note, I'll ask a fun and intriguing question whose answer I don't yet know:

**Question 2** Is there a way to generalize $S_n$ and $m(\cdot)$ so that the product $k \cdot \pi$ makes sense for some values of $k \in \mathbb{R}$, and so that $m(k \cdot \pi) = |k| \cdot m(\pi)$?

# References


Artin, M. 1991. *Algebra.* Prentice Hall.

Herstein, I. N. 1990. *Abstract Algebra.* Macmillan Pub.