1. Permutations or: group theory in 15 minutes

For those of you who already took a course in group theory, you probably learned about "abstract groups" which are sets with binary operations satisfying a list of conditions. I want to talk about permutations. Those who already know group theory can think about the question:

Why is every group isomorphic to a permutation group?

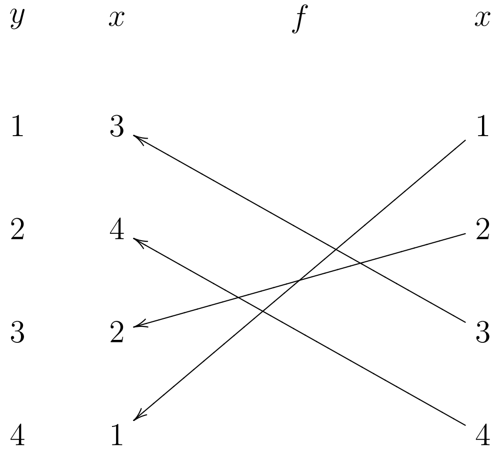
Definition 1.1. A *permutation* on a set X is a bijection $f : X \to X$. Recall that a *bijection* is a mapping which is

(1) 1-1 (injective) and
(2) onto (surjective).

If $X = \{1, 2, \cdots, n\}$, a permutation of X is called a *permutation on n letters*. The set of permutations of X will be denoted by $Perm(X)$.

I discussed three notations for permutations:

(1) $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 2 & 1 \end{pmatrix}$ means $\sigma(1) = 3, \sigma(2) = 4, \sigma(3) = 2, \sigma(4) = 1$ or $\sigma(x) = y$ where x is given in the first row and y is given in the second row.
(2) (cycle form) $\sigma = (1324)$. This means σ sends 1 to 3, 3 to 2, 2 to 4 and 4 to 1: $1 \to 3 \to 2 \to 4 \to 1$.
(3) (graphic notation). Here you put the numbers $x = 1, \cdots, n$ vertically on the right, put $y = 1, \cdots, n$ vertically on the left and connect each x to each $y = f(x)$. For example, if $f = (1423)$, you connect 1 on the right to $f(1) = 4$ on the left with a straight line, etc. In cycle notation, $f = (1423)$.



Definition 1.2. If $f, g \in Perm(X)$, $fg = f \circ g$ is the *composition* of f and g. This is the permutation defined by

$$fg(x) = f(g(x))$$

It means you do g first and then f.

Question: If $g = (34)$ what is gf? Write the answer in all three notations and demonstrate the composition in the notation.

(1) Stack up f, g, putting the first operation f on top and the second g underneath. Then cross out the second line:

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 1 & 2 \\ 3 & 4 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix}$$

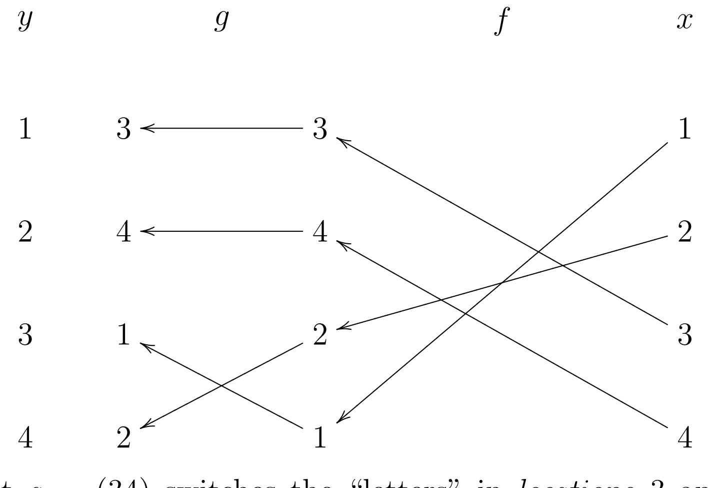
(2) Write the cycles next to each other and compute the image of each x by applying the cycles one at a time going from right to left:

$$gf = (34)(1423) = (13)(24)$$

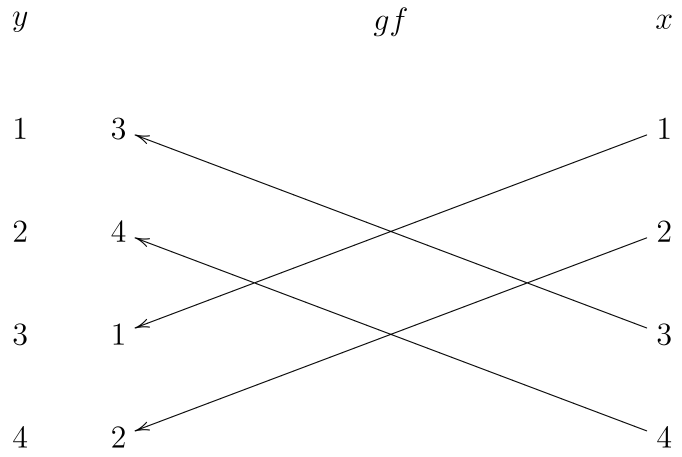
for example:

$$gf(1) =_3 (34)_4(1423)_1 = 3$$

(3) Draw the diagrams next to each other, putting the first permutation f on the right:



Note that $g = (34)$ switches the "letters" in *locations* 3 and 4. We discussed the fact that two of the crossings cancel when we redraw the picture:



Then we discussed the number of crossings.

2. Transpositions

On Day 2, I tried to formalize things we talked about at the end of Day 1, namely, in what way do the crossings in the diagram give transpositions and: What is the "longest word" ?

Definition 2.1. A *transposition* is a 2-cycle (ab). For example

$$(14) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 3 & 1 \end{pmatrix}$$

is a transposition. A *simple transposition* (also called a *simple reflection*) is a transposition of two consecutive letters:

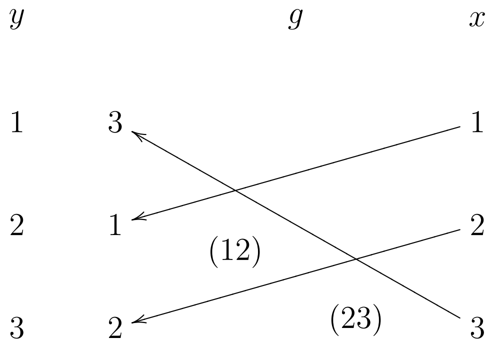
$$s_i := (i, i+1)$$

In the group of permutations of $n+1$ letters there are n simple reflections: $s_1 = (12), s_2 = (23), \cdots, s_n = (n, n+1)$.

We talked about the following vaguely phrased theorem.

Theorem 2.2. *A permutation f can be written as a product of simple reflections, one for every crossing in the diagram for f.*

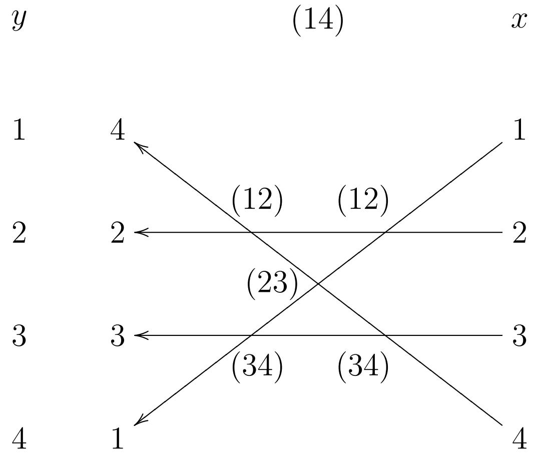
I drew a bunch of diagrams to illustrate this. The first was:



The crossing on the right is $s_2 = (23)$. The one on the left if $s_1 = (12)$. Reading these from right to left we get:

$$g = (123) = (12)(23) = s_1 s_2.$$

When two crossing lie above each other (so that you don't know which one is on the left and which is on the right) you can write them in either order. For example:



This gives:

$$(14) = (12)(34)(23)(12)(34) = s_1 s_3 s_2 s_1 s_3$$

But the crossings labeled $s_1 = (12)$ and $s_3 = (34)$ lie above each other. So we can write them in either order:

$$(14) = s_3 s_1 s_2 s_3 s_1$$

In general, the rule is

$$s_i s_j = s_j s_i \text{ if } |j - i| \geq 2$$

Definition 2.3. Suppose that f is a permutation of the letters $1, 2, \cdots, n+1$. Then the *length* $\ell(f)$ of f is defined to be the number of pairs of integers i, j so that

(1) $1 \leq i < j \leq n + 1$
(2) $f(i) > f(j)$.

In other words, it is the number of pairs of numbers whose order is switched by f.

Corollary 2.4. *Any permutation f can be written as a product of $\ell(f)$ simple reflections.*

Proof. Draw the diagram of the permutation. Move the lines slightly up or down so that there are no triple crossings of lines. Lines i and j will cross if and only if the are in one order on the right ($i < j$) and in the other order on the left ($f(i) > f(j)$). Therefore, the number of crossings is equal to $\ell(f)$ as defined above. The theorem says that each crossing gives one simple reflection and that f is a product of those simple reflections. Therefore, f is a product of $\ell(f)$ reflections. □

How large can $\ell(f)$ be?

The maximum possible number occurs when every single pair of integers $i < j$ gets reversed. This is the permutation called w_0:

$$w_0 = \begin{pmatrix} 1 & 2 & 3 & \cdots & n+1 \\ n+1 & n & n-1 & \cdots & 1 \end{pmatrix}$$

The length of w_0 is the number of all pairs $1 \leq i < j \leq n+1$ which is

$$\ell(w_0) = \binom{n+1}{2} = \frac{n(n+1)}{2} = 1 + 2 + \cdots + n$$

This is a *triangle number*. As Tri pointed out in class, w_0 can be written as a product of n s_1's, $n-1$ s_2's, etc. For example, for $n = 4$,

$$w_0 = \begin{array}{ccccccc} & & & s_4 & & & \\ & & s_3 & & s_3 & & \\ & s_2 & & s_2 & & s_2 & \\ s_1 & & s_1 & & s_1 & & s_1 \end{array}$$

$$= \quad s_1s_2s_1s_3s_2s_4s_1s_3s_2s_1$$

If we want to draw the diagram with 1 at the top and $n+1 = 5$ at the bottom, we could also write:

$$w_0 = \begin{array}{ccccccc} & & & s_1 & & & \\ & & s_2 & & s_2 & & \\ & s_3 & & s_3 & & s_3 & \\ s_4 & & s_4 & & s_4 & & s_4 \end{array}$$

In this method the simple reflection s_i occurs i times. And this is the method I used on Day 0. There are many different ways to write w_0 as a product of $\binom{n+1}{2}$ simple reflections. We will study this in detail later.

3. more group theory

Today we talked about the main properties of the elements of a group (although I didn't define a group yet). I just said that these are all permutations. I changed the format of the class to problem solving, similar to problem sessions, instead of lecturing.

3.1. inverse.

Definition 3.1. The inverse of a permutation σ is given by

$$\sigma^{-1}(x) = y \quad \text{where} \quad \sigma(y) = x$$

a) Show that this defines a permutation σ^{-1}.
b) Find a formula for the inverse of $\tau = (a_1, a_2, \cdots, a_k)$.
c) Show that $(\sigma\tau)^{-1} = \tau^{-1}\sigma^{-1}$.

Students found Question (a) confusing so we first did (b) and (c).

3.1.1. *inverse of a k-cycle.* The inverse of a cycle is given by writing the cycle backwards:

$$\tau^{-1} = (a_k, a_{k-1}, \cdots, a_2, a_1)$$

This is supposed to obvious, but a proof would go like this:

Proof. Let $x = a_i$. Then $\tau(a_{i-1}) = a_i = x$. So, $\tau^{-1}(a_i) = a_{i-1}$. Special care is needed in the case $i = 1$. Then the equation $\tau(a_k) = a_1$ means (by definition) that $\tau^{-1}(a_1) = a_k$. If x is not any of the a_i then $\tau(x) = x$. So, $\tau^{-1}(x) = x$. This shows that $\tau^{-1}(x)$ is given by the cycle above for all x. □

3.1.2. *inverse of a product.* The derivation of the formula

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

used a lot more stuff than I thought. The concepts that were used by students familiar with groups were the following:

(1) *associativity*: $(ab)c = a(bc)$. This implies that parentheses can be placed arbitrarily.
(2) *identity*: $id(x) = x$ is called the *identity* function and it is also called $e = id$. This has the property that $e\sigma = \sigma = \sigma e$ since:

$$e\sigma(x) = e(\sigma(x)) = \sigma(x), \quad \sigma e(x) = \sigma(e(x)) = \sigma(x)$$

(3) The group theoretic definition of the inverse which is:

$$\sigma^{-1}\sigma = e = \sigma\sigma^{-1}.$$

The proof of the inverse formula that students came up with, using these properties, was:

$$\begin{aligned}(\sigma\tau)(\tau^{-1}\sigma^{-1}) &= \sigma\underbrace{(\tau\tau^{-1})}_{e=id}\sigma^{-1} && \text{by associativity}\\ &= \sigma e\sigma^{-1} && \text{by group theoretic def of inverse}\\ &= \sigma\sigma^{-1} && \text{by property of } e\\ &= e && \text{by def of inverse}\end{aligned}$$

We needed to know that, when the product of two things is the identity, the two things are inverse to each other:

Lemma 3.2. *If $ab = e$ then $b = a^{-1}$.*

Proof. Multiply both sides by a^{-1}:

$$a^{-1}ab = a^{-1}e$$

The left hand side (LHS) is $a^{-1}ab = eb = b$, the RHS is a^{-1}. □

I actually skipped this lemma. What I verified in class was the group theoretic definition of inverse using my definition.

Lemma 3.3. *If σ^{-1} is defined as in Def. 3.1 (the "inverse function" definiton) then we get the formulas:*

$$\sigma\sigma^{-1}(x) = x\ \forall x,\ \text{i.e., } \sigma\sigma^{-1} = id$$

$$\sigma^{-1}\sigma(y) = y\ \forall y,\ \text{i.e., } \sigma^{-1}\sigma = id$$

Proof. My definition was that $\sigma^{-1}(x) = y$ if $\sigma(y) = x$. If we insert $\sigma(y)$ in for x in the first equation we get:

$$\sigma^{-1}(\sigma(y)) = y$$

Thus $\sigma^{-1}\sigma = id$. If we insert the first equation into the second we get:

$$x = \sigma(y) = \sigma(\sigma^{-1}(x))$$

i.e., $\sigma\sigma^{-1} = id$. □

3.1.3. *definition of a function.* I had to explain the first question because students had no idea even what it was asking.

Define: When I say that this formula *defines a function* I mean that for every x there is a unique y (so that $\sigma(y) = x$).

I used the fact that σ is a permutation ⇒ bijection ⇒ 1-1 and onto.

(1) (existence) y exists since σ is *onto*: This means for any $x \in X$ there is a $y \in X$ so that $\sigma(y) = x$. So, $\sigma^{-1}(x)$ exists.

(2) (uniqueness) We need to know that for each x there is only one y, otherwise we don't have a function. I gave the example of:

$$\sqrt{x} = y \quad \text{if } y^2 = x$$

This formula does not define the square root function since there are two y's for each positive x.

Uniqueness of y follows from the fact that σ is 1-1: If I had two y's say y_1 and y_2 (in other words, $\sigma^{-1}(x) = y_1$ and $\sigma^{-1}(x) = y_2$) then I would have $\sigma(y_1) = x = \sigma(y_2)$ which implies $y_1 = y_2$ since σ cannot send to y's to the same thing.

This show that $\sigma^{-1} : X \to X$ is a function. To show it is a permutation, we have to show it is 1-1 and onto. We decided in class that both statements are obvious when you write down what they are in equation form:

$$\sigma^{-1}(x_1) = y = \sigma^{-1}(x_2) \Rightarrow \sigma(y) = x_1, \sigma(y) = x_2$$

So, σ^{-1} is 1-1. To show it is onto, we need to take any y in our set X and find some x so that

$$\sigma^{-1}(x) = y, \quad \text{i.e.,} \quad \sigma(y) = x.$$

Well, there it is!

3.2. **commutativity.** Students brought up the subject of*commutativity.*

Definition 3.4. *a and b commute if $ab = ba$. A group G is commutative if $ab = ba$ for all $a, b \in G$.*

I pointed out that "commute" is a *verb* and "commutative" is an *adjective.*

Problem: If a, b commute, show that $(ab)^{-1} = a^{-1}b^{-1}$.

Proof. To show this you need to show

$$aba^{-1}b^{-1} = e$$

The proof that students gave is to switch a, b on the left since they commute:

$$aba^{-1}b^{-1} = baa^{-1}b^{-1} = bb^{-1} = e.$$

□

3.3. **conjugation.**

Definition 3.5. We say that a is *conjugate* to b if there exists a c so that

$$a = cbc^{-1}$$

I used the following notation. First, $a \sim b$ for "a is conjugate to b" and

$$\phi_c(b) := cbc^{-1}$$

The function ϕ_c is a "homomorphism" (a concept that I will explain later).

(1) Show that conjugation is an equivalence relation, i.e., it is
  (a) reflective: $(\forall a) a \sim a$
  (b) symmetric: $(\forall a, b) a \sim b \Rightarrow b \sim a$
  (c) transitive: $(\forall a, b, c) a \sim b, b \sim c \Rightarrow a \sim c$
(2) If $\tau = (a_1, a_2, \cdots, a_k)$ then find a formula for $\sigma\tau\sigma^{-1}$
(3) Show that every permutation of n letters is conjugate to its inverse.

3.3.1. *conjugacy is an equivalence relation.* We verified in class that the three properties of an equivalence relation hold:

(1) *reflexive*: $a \sim a$.
Just take $c = e$ (the identity). Then $a = eae^{-1} = \phi_e(a)$. So every permutation is conjugate to itself.
(2) *symmetry*: $a \sim b \Rightarrow b \sim a$.
We are given that $a = cbc^{-1}$. Multiply both sides on the left with c^{-1} and on the right with $c = (c^{-1})^{-1}$:

$$c^{-1}ac = c^{-1}a(c^{-1})^{-1} = c^{-1}\left(cbc^{-1}\right)c = ebe = b$$

So, $b = \phi_{c^{-1}}(a)$ and $b \sim a$.
(3) *transitive*: $a \sim b, b \sim x \Rightarrow a \sim x$. (We realized that the letter c was being used too often.)
We are given that $a = cbc^{-1}$ and $b = dxd^{-1}$. Insert this formula for b into the formula for a to get:

$$a = c(dxd^{-1})c^{-1} = (cd)x(d^{-1}c^{-1}) = (cd)x(cd)^{-1} = \phi_{cd}(x)$$

Here we used the formula $(cd)^{-1} = d^{-1}c^{-1}$ from earlier.

3.3.2. *conjugate of a k-cycle.* We found the formula and proved it.

Theorem 3.6. *If σ, τ are permutations and $\tau = (a_1, a_2, \cdots, a_k)$ then*

$$\sigma\tau\sigma^{-1} = (\sigma(a_1), \sigma(a_2), \cdots, \sigma(a_k)).$$

Proof. Take $x = \sigma(a_i)$. Then $\sigma^{-1}(x) = a_i$ by definition of the inverse σ^{-1} as discussed earlier. So,

$$\sigma\tau\sigma^{-1}(x) = \sigma\tau(a_i) = \sigma(a_{i+1})$$

This is for $i = 1, \cdots, k-1$. If $x = \sigma(a_k)$ then

$$\sigma\tau\sigma^{-1}(x) = \sigma\tau(a_k) = \sigma(a_1)$$

This almost does it. This shows that $\sigma\tau\sigma^{-1}$ moves the letters $\sigma(a_i)$ as indicated. But we also checked that the other letters are fixed. If x is not equal to any $\sigma(a_i)$ then $\sigma^{-1}(x)$ is not equal to any of the a_i which means that τ does not move it. So

$$\tau\left(\sigma^{-1}(x)\right) = \sigma^{-1}(x).$$

When you apply σ you get back x:

$$\sigma\tau\left(\sigma^{-1}(x)\right) = \sigma\sigma^{-1}(x) = x.$$

So, $\sigma\tau\sigma^{-1}$ fixes every letter not in the cycle $(\sigma(a_1), \cdots, \sigma(a_k))$ which means that

$$\sigma\tau\sigma^{-1} = (\sigma(a_1), \cdots, \sigma(a_k)).$$

□

3.3.3. *every permutation of n is conjugate to its inverse.* We did this by example. Suppose that τ is a cycle. For example $\tau = (123)$. Then

$$\tau^{-1} = (321) = (132) = (213)$$

By the conjugation formula above, we just need to choose σ which changes $1, 2, 3$ into the three letters in τ^{-1}. So, there are three answers:

(1) $\sigma = (13)$
(2) $\sigma = (23)$
(3) $\sigma = (12)$.

Each of these will conjugate τ into its inverse.

When there is more than one cycle we have to treat each cycle separately. For example:

$$\tau = (12)(345)(8, 9, 10, 11)$$

$$\tau^{-1} = (11, 10, 9, 8)(543)(12) = (12)(543)(11, 10, 9, 8)$$

Here I used two formulas. First,

$$(abc)^{-1} = c^{-1}b^{-1}a^{-1}$$

Then, I used the fact that disjoint cycles commute.

Theorem 3.7. *If σ, τ are permutations of disjoint sets of numbers then $\sigma\tau = \tau\sigma$.*

Then you do the same thing as before. For example let

$$\sigma = (45)(8,11)(9,10).$$

Then

$$\sigma\tau\sigma^{-1} = (\sigma(1)\sigma(2))(\sigma(3)\cdots \text{ etc.} = (12)(543)(11,10,9,8) = \tau^{-1}.$$

This is one of those proofs that are a real pain to write down in complete detail.

3.4. **definition of a group.** This spilled over to Day 4 (Wednesday).

Definition 3.8. A *group* is a set G with a binary operation (for all $a, b \in G$ there is a product $ab \in G$) satisfying three properties:

(1) (*associative*) $(ab)c = a(bc)$ for all $a, b, c \in G$.
(2) (*identity*) G has an identity element e so that $ae = ea = a$ for all $a \in G$. (Think: $e = 1$.)
(3) (*inverse*) Every element $x \in G$ has an inverse x^{-1} which is also an element of G so that

$$xx^{-1} = x^{-1}x = e.$$

The discussion from Monday proves the following.

Theorem 3.9. *Permutations of X forms a group, i.e., the set of all permutations of X is a group.*

Problem: Show that the set of 2×2 integer matrices with determinant ±1 forms a group under matrix multiplication. Thus

$$\det\begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc = \pm 1$$

and $a, b, c, d \in \mathbb{Z}$.

(1) (*associative*) We skipped the boring proof that matrix multiplication is associative.
(2) (*identity*) The identity is the matrix

$$I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

This satisfies the equation $I_2X = X = XI_2$ for any 2×2 matrix X.
(3) (*inverse*) The inverse of a 2×2 matrix A is

$$A^{-1} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

When $ad - bc = \pm 1$ this matrix also has integer entries. Liz pointed out that we also need to check that the determinant of the new matrix A^{-1} is ±1 because the definition of a group says the inverse needs to be an element of G. But this is a simple calculation:

$$\det \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \frac{1}{(ad-bc)^2}(ad-bc) = \frac{1}{ad-bc} = \frac{1}{\pm 1} = \pm 1$$

(Why does the scalar become squared in the determinant?)

I said that there is a general formula

$$\det(AB) = \det A \det B$$

So,

$$\det(AA^{-1}) = \det I_2 = 1 = \det A \det A^{-1}$$

which implies that

$$\det A^{-1} = \frac{1}{\det A}$$

3.5. **permutation matrices.**

Definition 3.10. If σ is a permutation of n letters, the *permutation matrix* $M(\sigma)$ is defined to be the matrix with entries p_{ij} where

$$p_{ij} = \begin{cases} 1 & \text{if } \sigma(j) = i \\ 0 & \text{if not} \end{cases}$$

For example, if $\sigma = (123)$ then

$$M(\sigma) = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

Questions:

(1) Show that $M(\sigma)M(\tau) = M(\sigma\tau)$
(2) Describe $M(\sigma)X$, i.e., what happens to X when you multiply on the left by $M(\sigma)$?
(3) What is $\det M(\sigma)$?

I think we did the last question first:

$$\det M(\sigma) = \pm 1.$$

You can prove this by induction. This will be homework.

To do question (2) we did the example of $\sigma = (123)$:

$$M(\sigma)X = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = \begin{pmatrix} g & h & i \\ a & b & c \\ d & e & f \end{pmatrix}$$

Thus multiplication by $M(\sigma)$ on the left moves the numbers in the first row into the second row, the second row moves to the third and the third row moves to the first.

Theorem 3.11. *Multiplication of a matrix X on the left by a permutation matrix $M(\sigma)$ permutes the rows of X by the permutation σ.*

Proof. Here is a formal proof. By definition of matrix multiplication,

$$(M(\sigma)X)_{ik} = \sum_{j=1}^{n} p_{ij} x_{jk}$$

But $p_{ij} = 1$ when $i = \sigma(j)$ or $j = \sigma^{-1}(i)$ and $p_{ij} = 0$ otherwise. This means that only one term in the sum is nonzero:

$$(M(\sigma)X)_{ik} = p_{ij} x_{jk} \quad \text{where } i = \sigma(j)$$

In other words, the (j, k) entry of X moves to the $(\sigma(j), k)$ position. Since this happens for every k, the entire j-th row of X moves to the $\sigma(j)$-th row. □

Students were surprised to figure out that the multiplication on the right has a different rule:

Theorem 3.12. *Multiplication of a matrix X on the right by a permutation matrix $M(\sigma)$ permutes the columns of X by the permutation σ^{-1}.*

With this rule we can figure out the proof for the 1st statement. Take a matrix X.

Lemma 3.13.

$$M(\sigma)M(\tau)X = M(\sigma\tau)X$$

Proof. By associativity, $(M(\sigma)M(\tau))X = M(\sigma)(M(\tau)X)$. But, by the previous discussion, this does the following. In $M(\tau)X$ the rows of X are permuted by τ. In $M(\sigma)(M(\tau)X)$ the rows are then permuted by σ. But, this is the permutation $\sigma\tau$ applied to the rows of X which is $M(\sigma\tau)X$. □

Theorem 3.14.

$$M(\sigma)M(\tau) = M(\sigma\tau)$$

Proof. Multiply by the inverse of X:

$$M(\sigma)M(\tau)\underbrace{XX^{-1}}_{I_n} = M(\sigma\tau)\underbrace{XX^{-1}}_{I_n}$$

$$M(\sigma)M(\tau) = M(\sigma\tau).$$

□