LITTLE TIDBITS ABOUT THE SYMMETRIC GROUP

Abstract. We give a very elementary introduction to the symmetric group. Section 1 gives the minimal background in symmetric groups. In sections 2 to 5 we discuss the Orbit-Stabilizer Theorem as well as Pólya theory of counting symmetries, applying it to problems such as enumerating vertex coloring of regular shapes and isomers of molecules. Section 6 proves Cayley's theorem for finite groups, section 7 talks about the Futurama theorem from the well-known animated series, and finally section 8 explains the math behind the 100 prisoners problem.

1. The symmetric group

The symmetric group is something that we are all familiar with, because it is simply the set of permutations of a finite set. Let us now make this precise. Note that we denote composition of two functions $f : A \longrightarrow B$ and $g : B \longrightarrow C$ as gf throughout instead of $g \circ f$.

Definition 1. Let $n \geq 1$ be an integer. The *symmetric group* $\mathfrak{S}_n$ is defined to be the set of all bijective functions on $\{1, \ldots, n\}$. An element of $\mathfrak{S}_n$ is called a *permutation.*

Proposition 2. *$\mathfrak{S}_n$ has the following properties:*

(a) it has cardinality $|\,\mathfrak{S}\,| = n!$,
(b) the composition of two permutations is a permutation,
(c) composition of permutations is associative, i.e. that $(fg)h = f(gh)$ for all $f, g, h \in \mathfrak{S}_n$,
(d) the identity permutation $1_n \in \mathfrak{S}_n$, defined by $1_n(k) = k$ for all $k \in \{1, \ldots, n\}$, has $1_n f = f 1_n = f$ for all $f \in \mathfrak{S}_n$,
(e) every $f \in \mathfrak{S}_n$ has a unique inverse function $f^{-1} \in \mathfrak{S}_n$ such that $ff^{-1} = f^{-1}f = 1_n$.
(f) If $f, g \in \mathfrak{S}_n$, then $(fg)^{-1} = g^{-1}f^{-1}$.

Proof. Easy consequence of the basic property of functions, and left to the reader. □

The most common way to represent elements of $\mathfrak{S}_n$ is by a 2 by n matrix, where the first row is simply writing down 1 to n in the usual integer ordering, and the second row is writing down the entries after permuting it. For example, in $\mathfrak{S}_3$, the element that permutes 1 and 3 and leaves 2 fixed is given by

$$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix},$$

while the element right translating $1, 2, 3$ by one is given by

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}.$$

Note that it is quite a pain to write two rows, so we can represent them by something called *cycle factorization*, where we write elements into *disjoint cycles*. Every disjoint cycle of length k is called a *k-cycle.* It's best to explain these in terms of an example. Consider the element

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 2 & 4 & 8 & 1 & 5 & 3 & 6 & 7 \end{pmatrix} \in \mathfrak{S}_8\,.$$

Elements that cycles in a loop are grouped together and written as a cycle. Here we see 1 maps to 2, which maps to 4, which maps back to 1, so we can write them as (1 2 4). Continuing in this manner we see the cycle factorization of σ is $\sigma = (1\ 2\ 4)(3\ 8\ 7\ 6)(5)$. It is a convention to drop all 1-cycles in a cycle factorization, so we can also write

$$\sigma = (1\ 2\ 4)(3\ 8\ 7\ 6).$$

Since σ is seen to be a function, we can perform our usual function composition and function evaluation *from right to left.* For example, $\sigma(1) = 2$ and $\sigma^{-1}(6) = 7$ and $\sigma(5) = 5 = \sigma^{-1}(5)$. If we let

$$\tau = (1\ 3\ 5)(2\ 8)(4\ 6\ 7),$$

then $(\tau\sigma)(1) = \tau(2) = 8$ and $(\tau\sigma)^{-1}(2) = (\sigma^{-1}\tau^{-1})(2) = \sigma^{-1}(8) = 3$.

We now study a geometrical subset of $\mathfrak{S}_n$, called the dihedral group. One way to describe it is the set of all *symmetries* of a regular n-gon, i.e. all bijections of vertices of the n-gon that preserves the distance between each vertex. For example, reflections and rotations about an axis are all symmetries of a regular n-gon. In fact, the proposition below tells us that these are all the symmetries on a regular n-gon.

Proposition 3. *There are exactly $2n$ symmetries of a regular n-gon.*

Proof. Consider two adjacent vertices v_1 and v_2 on the regular n-gon. We can send v_1 to the n possible positions of the vertices of the regular n-gon. Then v_2 has two possible positions, either to the left or right of v_1. This gives $2n$ possible symmetries, and it's clear the symmetries are pairwise distinct. □

There is an abstract way to match every symmetry of a regular n-gon into an element of $\mathfrak{S}_n$. By labeling the vertices of the regular n-gon counterclockwise by the numbers $\{1, \ldots, n\}$, the symmetries are:

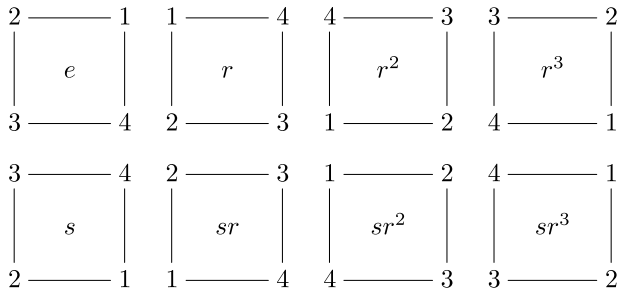
- the set of n rotations $\mathfrak{R}_n = \{e, r, \ldots, r^{n-1}\}$, where we define $r := (1\ \ 2\ \ \cdots\ \ n)$,
- the set of n reflections $\mathfrak{r}_n = \{s, sr, \ldots, sr^{n-1}\}$, where we define

$$s := \begin{cases} (1\ \ n)(2\ \ n-1)\cdots(\frac{1}{2}(n-2)\ \ \frac{1}{2}(n+2)) & \text{if } n \text{ is even,} \\ (2\ \ n)(3\ \ n-1)\cdots(\frac{1}{2}(n-1)\ \ \frac{1}{2}(n+3)) & \text{if } n \text{ is odd.} \end{cases}$$

Geometrically, the r corresponds to rotation by $2\pi/n$ radians counterclockwise about the center of the n-gon, and the s corresponds to reflection about some axis of symmetry depending on the parity of n.

Definition 4. The *dihedral group* D_n is the subset of $\mathfrak{S}_n$ with elements $\mathfrak{R}_n \sqcup \mathfrak{r}_n$.

Example 5. D_4 is shown in terms of pictures as the eight symmetries of a square below.



The following proposition tells us r and s behaves like what we want. Let us define the *order* of $g \in \mathfrak{S}_n$ to be

$$|g| := \begin{cases} \min\{n \in \mathbb{Z}_{>0} : g^n = e\} & \text{if there exists } n \in \mathbb{Z}_{>0} \text{ with } g^n = 1_n, \\ \infty & \text{otherwise.} \end{cases}$$

Proposition 6. *With the notation of an order of an element in $\mathfrak{S}_n$ as above, $|s| = 2$, $|r| = n$, and $rs = sr^{-1}$.*

Proof. Left as an easy computational exercise using the definitions of r and s above. □

Let us now show that the abstract construction of the symmetries of a regular n-gon satisfies basic properties analogous to those listed in proposition 2. The proof is very combinatorial, as is most arguments of basic group theory.

Proposition 7. *D_{2n} has the following properties:*

(a) it has cardinality $2n$,
(b) if $f, g \in D_{2n}$, then $gf \in D_{2n}$,
(c) composing elements in D_{2n} is associative,
(d) the identity permutation 1_n is in D_{2n},
(e) every $f \in D_n$ has a unique inverse $f^{-1} \in D_n$ such that $ff^{-1} = f^{-1}f = 1_n$.

Proof. D_{2n} has cardinality $2n$ by proposition 3. Associativity follows because $D_{2n} \subset \mathfrak{S}_n$, and the identity is in D_{2n} by definition. For inverses, r^{n-k} is the inverse to r^k and sr^k is its own inverse for $k \in \{1, \ldots, n\}$, as

- $r^k r^{n-k} = r^{k+(n-k)} = r^n = e = r^{(n-k)+k} = r^{n-k}r^k$,
- $(sr^k)(sr^k) = s(r^{k-1}sr^{-1})r^k = s(r^{k-2}sr^{-2})r^k = \cdots = s(sr^{-k})r^k = s^2e = e$, using the identities $rs = sr^{-1}$ and $|s| = 2$.

Finally to show part (b). Let $s^i r^j, s^x r^y \in D_{2n}$ for $i, x \in \{0, 1\}$ and $j, y \in \{1, \ldots, n\}$. Consider the following procedure:

- Choose the rightmost s in α, where $\alpha := s^i r^j s^x r^y$.
- If this s is at the right of any other s, multiply them so that they equal the identity and we lose two terms in $a_1 a_2 \cdots a_k$. Otherwise an r is at the left of this s, so use the relation $rs = sr^{-1}$ in proposition 6 to push this s left.
- Repeat the procedure until there is no s at the right of any r in α.

In the end we will get that $\alpha = s^a r^b$ with $a \in \{0, 1\}$ and $b \in \{1, \ldots, n\}$ after using the fact that s has order 2 and r has order n, and so $\alpha \in D_{2n}$. □

Exercise 8. *Show that $\mathfrak{R}_n$ also satisfy properties (b) to (e) in proposition 7, but $\mathfrak{r}_n$ does not satisfy properties (b) and (c).*

We introduce a final important subset of $\mathfrak{S}_n$ called the alternating group.

Definition 9. Let $n \geq 2$. The *alternating group* is the subset of $\mathfrak{S}_n$ containing the set of elements which can be written as an even number of compositions $s_1 \cdots s_{2k}$, where $s_1, \ldots, s_{2k} \in \mathcal{S} := \{(a\ \ b) : 1 \leq a < b \leq n\}$. Elements in the set $\mathcal{S}$ are called *transpositions* of $\mathfrak{S}_n$.

We want to show A_n has similar properties as $\mathfrak{S}_n$ and D_{2n}. But this is a little trickier than the above. In particular we will need to prove two lemmas first.

Lemma 10. *For $n \geq 2$, every element in $\mathfrak{S}_n$ is a product of transpositions.*

Proof. As we can write every element $\sigma \in \mathfrak{S}_n$ as a composition of disjoint cycles, it suffices to show this for some $(1\ \ 2\ \ \cdots\ \ k) \in \mathfrak{S}_n$ after relabeling. But observe $(1\ \ 2\ \ \cdots\ \ k) = (1\ \ k)(1\ \ k-1)\cdots(1\ \ 2)$. □

Exercise 11. *In fact, every element in $\mathfrak{S}_n$ can be written as a composition of the transpositions from $\{(1\ \ k) : k \in \{1, \ldots, n\}\}$ since $(a\ \ b) = (1\ \ a)(1\ \ b)(1\ \ a)$. Strengthen lemma 10 in two more different directions as below by doing similar computations.*

(a) Every element in $\mathfrak{S}_n$ can be written as a composition of elements from $\{(1\ \ 2), (2, 3), \ldots, (n-1\ \ n)\}$.

(b) Every element in $\mathfrak{S}_n$ can be written as a composition of elements from $\{(1\ \ 2\ \ \cdots\ \ n-1\ \ n), (1\ \ 2)\}$.

Lemma 12. *For any $\sigma \in \mathfrak{S}_n$ with $n \geq 2$, let $\operatorname{sgn}(\sigma)$ be the number of elements modulo 2 after writing σ as a product of transpositions. Then $\operatorname{sgn}(\sigma)$ is well-defined, i.e. every element in $\mathfrak{S}_n$ can be written either as an even or odd number of transpositions, but not both.*

Proof. It suffices to show the lemma for the identity permutation $1 := 1_n$. To see this, suppose $\tau_1 \cdots \tau_n$ and $\tau'_1 \cdots \tau'_m$ are two ways to write σ as transpositions. Then $1 = \tau'_1 \cdots \tau'_m \tau_n^{-1} \cdots \tau_1^{-1}$. If $\operatorname{sgn}(1)$ is well-defined, this shows $m + n \equiv 0 \pmod 2$, so $m \equiv n \pmod 2$.

Let us now show $\operatorname{sgn}(1)$ is well-defined. We must show 1 always is an even product of transpositions. Clearly 1 cannot be written as a transposition, and for any transposition $(a_1\ \ a_2)$ with $a_1 \neq a_2$, we can write $1 = (a_1\ \ a_2)(a_1\ \ a_2)$. Now by induction suppose for some $k \geq 2$ that any way of writing 1 as a product of less than k transpositions must be even.

Say $\tau = (a_1\ \ b_1)\cdots(a_k\ \ b_k)$ equals 1, with some of the a_i and b_i possibly equal. Note that disjoint cycles commute (which means $\sigma\tau = \tau\sigma$ if τ and σ are disjoint cycles) and $(c_i\ \ c_j)(c_i\ \ c_k) = (c_k\ \ c_j)(c_i\ \ c_j)$. Hence, by picking the leftmost transposition of τ containing a_1 other than $(a_1\ \ b_1)$, we can perform this until the leftmost two transpositions are either of the form $(a_1\ \ b_1)(a_1\ \ b_1) = 1$ or $(a_1\ \ b_1)(a_1\ \ b_2) = (a_1\ \ b_2)(b_1\ \ b_2)$.

In the first case we now have $k - 2$ transpositions, which must be even by induction, so k is even. For the second case, iterate the procedure in the above paragraph, which must only be finitely many times as we only have finitely many transpositions with a_1. If we always land in the second case, then in the end we will get the leftmost transposition is the only one containing a_1, so that $1(a_1) = b_1 \neq a_1$, a contradiction to 1 being the identity of $\mathfrak{S}_n$. Hence we will eventually land in the first case after a finite number of steps, so we can finish by induction. □

We are finally ready to prove the following proposition.

Proposition 13. *A_n has the following properties:*
(a) it has cardinality $n!/2$,
(b) if $f, g \in A_n$, then $gf \in A_n$,
(c) composing elements in A_n is associative, that is, $(fg)h = f(gh)$ for all $f, g, h \in A_n$,
(d) the identity permutation 1_n is in A_n,
(e) every $f \in A_n$ has a unique inverse $f^{-1} \in A_n$ such that $ff^{-1} = f^{-1}f = 1_n$.

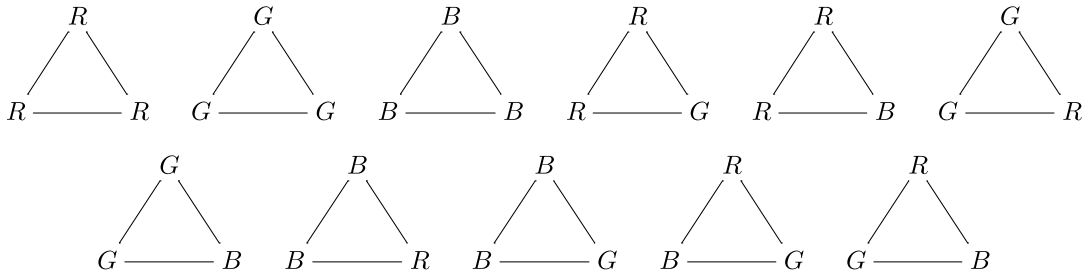
Proof. The last four properties are clear by the well-definedness of sgn. Let us show A_n is a group of cardinality $n!/2$. We demonstrate a bijection between it and the set O_n of elements which can be written as an odd number of transpositions. The bijection we seek is $f : A_n \longrightarrow O_n$ by $f(\tau) = (1\ 2)\tau$, which is again well-defined by the well-definedness of sgn. It is a bijection because the inverse is given by $f^{-1} = f$. This would imply what we want as $A_n \cup O_n = \mathfrak{S}_n$ by lemma 10 and $A_n \cap O_n = \emptyset$ by lemma 12, so that $|A_n| = |O_n|$ and A_n has cardinality $|\,\mathfrak{S}_n\,|/2 = n!/2$. □

Generally we do not like to consider the set O_n described in the proof of proposition 13. This is because composing two transpositions lands us inside A_n, not O_n.

2. Burnside's lemma: applications

In this section we state the symmetric group version of Burnside's lemma and give various applications of it, deferring the proof of the general statement to the next section. Let us first give a motivational example of the type of questions Burnside's lemma will be useful in.

Example 14. Let us look at how many nonequivalent ways there are to color the vertices of the triangle with colors R, G, B *under rotation*. An easy listing tells us there are 11 of them as follows.



However, there are only 10 nonequivalent colorings *under both rotation and reflection*, because the last two colorings listed above are equivalent under a reflection.

Certainly we can enumerate other symmetries similarly, but they are not easy to do directly if we have a large data. Hence we will discuss a tool called Burnside's lemma. In fact this is the key step to Pólya's enumeration theorem, which is the main theorem of Pólya theory and discussed in section 4. Note that we can already use Burnside's lemma to compute lots of stuff, and it is easier to compute things using this than Pólya's enumeration theorem (of course some data is lost and we can't get more interesting results). As promised let us now give the symmetric group version of Burnside's lemma. We need two preliminary definitions.

Definition 15. A subset G of the symmetric group $\mathfrak{S}_n$ is a *permutation group* if it satisfies the following four properties under function composition:
- if $f, g \in G$, then $fg \in G$,
- composition of permutations in G is *associative*, i.e. that $(fg)h = f(gh)$ for all $f, g, h \in G$,
- G contains the identity permutation 1_n, defined by $1_n(k) = k$ for all $k \in \{1, \ldots, n\}$,
- every $f \in G$ has a unique inverse function $f^{-1} \in G$ such that $ff^{-1} = f^{-1}f = 1_n$.

Example 16. The *trivial group* $\{1_n\}$ is a permutation group. The sets $\mathfrak{R}_n$ and D_n and A_n discussed in the previous section are permutation groups, but not $\mathfrak{r}_n$.

Definition 17. Let G be a permutation group, and let S be a set. We say G *acts* on S if there is a map $G \times S \longrightarrow S$ by $(g, h) \mapsto g \cdot h$ such that the following holds:

- $gh \cdot s = g \cdot (h \cdot s)$ for all $g, h \in G$ and $s \in S$,
- $1_G \cdot s = s$ for all $s \in S$.

We also define the following notions:

- the *orbit* of an element $s \in S$ is $\mathcal{O}_s := \{g \cdot s : g \in G\}$,
- the *stabilizer* of an element $s \in S$ is $\mathcal{G}_s := \{g \in G : g \cdot s = s\}$,
- the *fixed set* of an element $g \in G$ is $\mathrm{Fix}(g) := \{s \in S : g \cdot s = s\}$.

Example 18. The symmetric group $\mathfrak{S}_n$ (or any permutation group) acts on any set S with n elements naturally as follow: writing $S = \{s_1, \ldots, s_n\}$, the group action is defined to be the action on indices, that is, $\sigma \cdot s_i = s_{\sigma(i)}$. The orbit $\mathcal{O}_{s_i} = S$ and the stabilizer $\mathcal{G}_s = \{\sigma \in \mathfrak{S}_n : \sigma(i) = i\}$. The fixed points of any $\sigma \in \mathfrak{S}_n$ is $\mathrm{Fix}(\sigma) = \{s_j \in S : \sigma(j) = j\}$.

Finally, the symmetric group version of Burnside's lemma, followed by applications.

Proposition 19 (Burnside, symmetric group version)**.** *Let G be a permutation group acting on a set S, and let S/G be the number of distinct orbits. Then the cardinality of S/G can be computed as*

$$|S/G| = \frac{1}{|G|} \sum_{g \in G} |\operatorname{Fix}(g)|.$$

To make sure the right hand side is always defined we need $|G| \geq 1$. But this is true since by the definition of a permutation group G contains the identity permutation 1_n.

Example 20 (Coloring of vertices of regular n-gons)**.** Proposition 19 doesn't seem to relate to colorings of vertices of regular n-gons, but in fact it does. We can consider the permutation group D_n, and let S be the set of all k^n possible coloring of the vertices of the regular n-gons with k colors. Then we can define an action of D_n on S simply by symmetries, so that, for some $g \in D_n$:

- the orbit $\mathcal{O}_g$ is the set of all colorings of an n-gon that are pairwise related by a rotation or a reflection,
- the fixed set $\mathrm{Fix}(g)$ is the set of all colorings $s \in S$ of an n-gon that are *preserved* under the action of g, i.e. that we cannot tell s and $g \cdots s$ apart since $s = g \cdot s$.

Note that we can also look at the different types of colorings on the regular n-gon, for example on the edges, but we will not consider it here since they all use similar steps.

Now for a concrete computation of this example. Consider the square (regular 4-gon), and suppose we have 3 colors. All symmetries of the square (regular 4-gon) are listed in example 5. Denote S to be all $3^4 = 81$ colorings of the vertices of the square. We demonstrate how to count $\mathrm{Fix}(x)$ for each $x \in D_4$ using the two cases below.

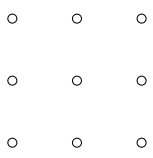
- If we take $r \in D_{2n}$, then all colorings of the square fixed under the action of r are those where all vertices are colored with the same color, so $|\operatorname{Fix}(r)| = 3$.
- If we take $sr \in D_{2n}$, then all colorings of the square fixed under the action of sr are those with vertices 1 and 3 the same color, and vertices 2 and 4 can be colored independently, so $|\operatorname{Fix}(sr)| = 3^3 = 27$.

Continuing in the same manner for the remaining six elements in D_{2n} and using the fact that $|D_4| = 8$, we see there are 21 nonequivalent colorings here. In fact, by a similar proof, given k colors we have

$$|S/D_4| = \frac{k^4 + 2k^3 + 3k^2 + 2k}{8}.$$

Although a bit overkill, this also give a combinatorial proof that $(k^4 + 2k^3 + 3k^2 + 2k)/8$ is always integer-valued if k is a positive integer.

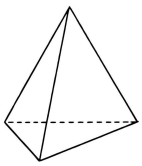
Exercise 21. *Suppose we have $k \geq 1$ colors. Find the number of different ways to color the nine identical points (under dihedral group action) on a square with k colors as shown.*



Exercise 22. *Find a general equation to compute the number of ways to color a regular n-gon with k colors. A way to do this is to consider the set of rotations $\mathfrak{R}_n$ and the set of reflections $\mathfrak{r}_n$ separately.*

Burnside's lemma can also be used to count symmetries of 3-dimensional objects *under rotations* (which is clearly still a group) in the same sense as for polygons. We do not consider three-dimensional reflections in this case because generally we cannot do that in the real world. If we are more chemistry-minded, this is because we want to avoid chiral compounds.

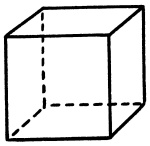
Example 23 (Rotation group of the tetrahedron)**.** Consider the tetrahedron, which is the regular polygon with 4 triangular faces, 6 edges, and 4 vertices as below.



Let us show the set of rotations R of the tetrahedron is *isomorphic* to the alternating group A_4, i.e. there is a bijection $f : R \longrightarrow A_4$ such that $f(r_1 r_2) = f(r_1)f(r_2)$ for any $r_1, r_2 \in R$, where we can compose two rotations in the usual way. Notice a 2-cycle corresponds to a reflection (not a rotation!) through an edge of the tetrahedron. Also notice a 3-cycle corresponds to a rotation through a vertex and the center of the opposite face. It can be easily seen a 4-cycle does not correspond to a rotation. A composition of two disjoint transpositions is also a rotation through the center of two opposite edges. Hence by lemma 10 the rotation group of the tetrahedron is indeed A_4, and we can construct an isomorphism $f : R \longrightarrow A_4$ just by assigning some labeling $1, 2, 3, 4$ to the vertices.

With this we can compute how many ways to color the four vertices of the tetrahedron with n colors just as in example 20, which turns out to be equal to $(n^6 + 8n^2 + 3n^4)/12$. The only difference is that one will now have $12 = |A_4|$ cases to compute out instead of 8 in example 20.

Exercise 24 (Symmetries of the cube)**.** *Suppose we have $k \geq 1$ colors. Let us consider the cube, which is the regular polytope with 6 square faces, 12 edges, and 8 vertices as below.*



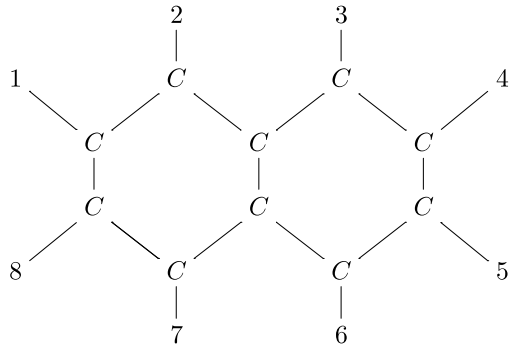
(a) Show that the set of rotations R of the cube is isomorphic to the symmetric group S_4, i.e. that there is a bijection $f : R \longrightarrow A_4$ such that $f(r_1 r_2) = f(r_1)f(r_2)$ for any $r_1, r_2 \in R$, where we can compose two rotations in the usual way.

(b) Show that there are $(n^6 + 3n^4 + 12n^3 + 8n^2)/24$ distinct ways to color the six faces of the cube with n colors under rotation.

(c) Find the number of distinct ways to color the faces of a cube (under rotation) with k colors such that two opposite faces have the same colors. Note that there are three sets of opposite faces.

(d) Suppose Eric and Kyle colors two cube with the requirements given in in part (c). Pick j rotations of the 24 at random, with $1 \leq j \leq 24$. Find the probability both their colorings are the same and each rigid motion of the j fixes at least f pairs of faces, with $0 \leq f \leq 3$.

Example 25 (Enumerating isomers of tetramethylnaphthalene)**.** Time for a chemistry application. The hydrocarbon naphthalene has ten carbon atoms (C) arranged in a double hexagon, and eight hydrogen atoms (H) attached at each of the corners of the hexagons. Tetramethylnaphthalene is obtained by replacing four of the hydrogen atoms of naphthalene with methyl groups (CH_3). We want to find how many *isomers* there are. That is, we want to find the number of ways to label the numbers in the following diagram with four H and four CH_3.



First up is to find the symmetry group of the diagram above (under rotation in three-dimensional space). It is easily seen to be the subset

$$V := \{1, (1\ 5)(2\ 6)(3\ 7)(4\ 8), (1\ 4)(2\ 7)(3\ 6)(5\ 8), (1\ 8)(2\ 3)(4\ 5)(6\ 7)\}$$

of S_8, noting that V is indeed a permutation group (the reader who knows some algebra will notice this is isomorphic to the *Klein 4-group*). Next, we notice that

$$|\operatorname{Fix}(1)| = \binom{8}{4} = 70,$$
$$|\operatorname{Fix}(x)| = \binom{4}{2} = 6 \text{ for each } x \in V \setminus \{1\}.$$

The first equality is the computation for all possible arrangements not considering symmetries, which is equivalent to asking the number of ways to arrange four H and four CH_3 in a line. The second equality is equivalent to asking the number of ways to arrange two HH and two CH_3CH_3 in a line. Hence proposition 19 tells us that the number of isomers of tetramethylnaphthalene is $(70 + 3 \cdot 6)/4 = 22$.

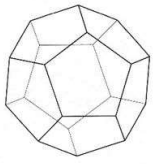
Exercise 26. *A Möbius strip is a 3-dimensional figure with only one suface. To get a Möbius strip, get a strip of paper, twist it by 180 degrees along two opposite edges and glue these two edges together. A picture is given below.*

Here we want to discuss the ways to color a Möbius strip by dividing it into n equal pieces on its surface by drawing $n \geq 1$ lines perpendicular to its edges.

a) Find all rotations of the Möbius strip for a fixed n.

b) Find the number of different ways to color the n pieces on its surface with $k \geq 1$ colors.

Exercise 27. *Consider the dodecahedron, which is the regular polygon with 12 pentagonal faces, 30 edges, and 20 vertices as below.*



Show that there are 9099 distinct ways to color the faces of a dodecahedron with three colors up to rotation.

3. Burnside's lemma: proof

In the previous two sections we considered some basic definitions of symmetric groups that can be generalized nicely to a topic known as *group theory*. We now generalize the definitions given in definitions 15 and 17 to the group theory setting.

Definition 28. Let G be a set. A *binary operation* on G is a function $* : G \times G \longrightarrow G$, and we write $g * h := *(g, h)$ for $g, h \in G$ (we write write gh to mean $g * h$ as well). A *group* is a set G with a binary operation $*$ satisfying the following properties:

- *associativity*, i.e. $(f * g) * h = f * (g * h)$ for all $f, g, h \in G$,
- existence of an *identity* $1 \in G$ such that $1 * g = g * 1 = g$ for all $g \in G$,
- every element $g \in G$ has an *inverse* $g^{-1} \in G$ such that $g * g^{-1} = g^{-1}g = 1$.

A subset H of G such that $x * y \in H$ for all $x, y \in H$ and satisfying the above three properties is called a *subgroup* of H.

Exercise 29. *Let G be a group. Show that the identity of G is unique, and the inverse of every element is unique. Also give an example to show that in general $gh \neq hg$ for some $g, h \in G$.*

Definition 30. Let G be a group, and let S be a set. We say G *acts* on S if there is a map $G \times S \longrightarrow S$ by $(g, h) \mapsto g \cdot h$ such that the following holds:

- $gh \cdot s = g \cdot (h \cdot s)$ for all $g, h \in G$ and $s \in S$,
- $1_G \cdot s = s$ for all $s \in S$.

We also define the following notions:

- the *orbit* of an element $s \in S$ is $\mathcal{O}_s := \{g \cdot s : g \in G\}$,
- the *stabilizer* of an element $s \in S$ is $\mathcal{G}_s := \{g \in G : g \cdot s = s\}$,
- the *fixed set* of an element $g \in G$ is $\mathrm{Fix}(g) := \{s \in S : g \cdot s = s\}$.

Proposition 31. *Let G be a group acting on a set S. Then, by restricting the binary operation of G to $\mathcal{G}_s$ for any $s \in S$, $\mathcal{G}_s$ is a subgroup of G.*

Proof. Let $x, y \in \mathcal{G}_s$. Then $x \cdot s = s$ and $h \cdot s = s$. Hence $xy \in \mathcal{G}_s$ since $xy \cdot s = x \cdot (y \cdot s) = x \cdot s = s$. Associativity is direct as $\mathcal{G}_s$ is a subset of G, and $1 \in G$ since $1 \cdot s = s$. Finally, if $g \in G$, then $g \cdot s = s$, so $g^{-1} \in G$ as $g^{-1} \cdot s = g^{-1} \cdot (g \cdot s) = g^{-1}g \cdot s = s$. □

We will now show that, for any group G acting on a set S, the orbits of S actually partition S, i.e. that $\bigcup_{s \in S} \mathcal{O}_s = S$, and that for every $s, t \in S$ either $\mathcal{O}_s \cap \mathcal{O}_t = \emptyset$ or $\mathcal{O}_s = \mathcal{O}_t$. To do this it is convenient to introduce the notation of an equivalence relation.

Definition 32. Let X be a set. We say $\sim$ is a *relation* on X if it is a subset of $X \times X$. If $(x, y) \in X \times X$, we say that x is *related* to y, and in notation $x \sim y$. We also say $\sim$ is an *equivalence relation* on X if it is a relation that is:

- *reflexive*, i.e. $x \sim x$ for all $x \in X$,
- *symmetric*, i.e. for all $x, y \in X$, if $x \sim y$, then $y \sim x$,
- *transitive*, i.e. for all $x, y, z \in X$, if $x \sim y$ and $y \sim z$, then $x \sim z$.

Also define $[x] := \{y \in X : y \sim x\}$ to be the *equivalence class* of an element $x \in X$.

The next proposition tells us equivalence relations are good because it partitions a set.

Proposition 33. *The equivalence classes of X with respect to an equivalence relation $\sim$ partitions X. That is, for $x, y \in X$, necessarily $[x] = [y]$ if $x \sim y$ and $[x] \cap [y] = \emptyset$ if $x \not\sim y$.*

Proof. Suppose $x \sim y$ with $x, y \in X$. We want to show $[x] = [y]$. Let $z \in [x]$. Then $z \sim x$. But $x \sim y$, so by transitivity $z \sim y$, implying $z \in [y]$ and $[x] \subseteq [y]$. Similarly letting $z \in [y]$, then $z \sim y$. But $x \sim y$, so by symmetry $y \sim x$ and by transitivity $z \sim x$, implying $z \in [x]$ and $[y] \subseteq [x]$. Thus $[x] = [y]$.

Now suppose $x \not\sim y$ with $x, y \in X$. Also suppose that $[x] \cap [y] \neq \emptyset$. Then there exists $z \in [x] \cap [y]$, so that $z \sim x$ and $z \sim y$. By symmetry $x \sim z$, and by transitivity $x \sim y$, a contradiction. Thus $[x] \cap [y] = \emptyset$. □

Using this language of equivalence relations, it suffices to show that the relation of being in the same orbit of a group action is an equivalence relation to show that orbits partition the set.

Proposition 34. *Let G be a group acting on a set S. Define a relation $\sim$ on S by $a \sim b$ if $b = g \cdot a$ for some $g \in G$. Then $\sim$ is an equivalence relation on S. In particular, $|S| = \sum_{s \in S} |\mathcal{O}_s|$ by proposition 33.*

Proof. Note that $1 \cdot s = s$ for all $s \in S$, so that $s \sim s$ and $\sim$ is reflexive. To show symmetry, suppose $s, t \in S$ such that $s \sim t$. Then $g \cdot s = t$ for some $g \in G$. Then $g^{-1} \cdot t = g^{-1} \cdot (g \cdot s) = g^{-1}g \cdot s = 1 \cdot s = s$, so that $t \sim s$. Finally to show transitivity, suppose $s, t, u \in S$ such that $s \sim t$ and $t \sim u$. Then $g \cdot s = t$ and $h \cdot t = u$ for some $g, h \in G$. Then $hg \in G$ with $hg \cdot s = h \cdot (g \cdot s) = h \cdot t = u$, so that $s \sim u$. □

To prepare for Burnside's lemma we will need an important result.

Lemma 35 (Orbit-Stabilizer)**.** *Let G be a finite group acting on S, and let $s \in S$. Then $|G| = |\mathcal{O}_s||\mathcal{G}_s|$. Hence, if $s, t \in S$ are in the same orbit, then $|\mathcal{G}_s| = |\mathcal{G}_t|$.*

Proof. We prove the first statement as the second statement is a direct consequence of the first. There is a clear surjection $f : G \longrightarrow \mathcal{O}_s$ by $f(g) = g \cdot s$. Note that $\{f^{-1}(x) : x \in \mathcal{O}_s\}$ partitions G. To see this, if $x = y$ in S then certainly $f^{-1}(x) = f^{-1}(y)$, and if $f^{-1}(x) \cap f^{-1}(y) \neq \emptyset$ then there exists $g \in G$ so that $x = g \cdot s = y$. To finish we need to show that $|f^{-1}(x)| = |f^{-1}(y)|$ for all $x, y \in S$, so that $|G| = \sum_{x \in \mathcal{O}_s} |f^{-1}(x)| = |\mathcal{O}_s||f^{-1}\{s\}| = |\mathcal{O}_s||\mathcal{G}_s|$.

Let us now show that $|f^{-1}(x)| = |f^{-1}(y)|$ for all $x, y \in S$. We see that $|f^{-1}(g \cdot s)| = |\mathcal{G}_{g \cdot s}|$ since there is the bijection $\varphi : f^{-1}(g \cdot s) \longrightarrow \mathcal{G}_{g \cdot s}$ by $\varphi(h) = g^{-1}h$ with inverse $\varphi^{-1}(h) = gh$. Note that φ is well-defined, since if $h \in f^{-1}(g \cdot s)$ then $h \cdot s = g \cdot s$, so that $\varphi(h) \cdot s = g^{-1}h \cdot s = s$. Hence it suffices to prove $|\mathcal{G}_{g \cdot s}| = |\mathcal{G}_s|$ for each $g \in G$, which is true since there is the bijection $\phi : \mathcal{G}_{g \cdot s} \longrightarrow \mathcal{G}_s$ by $\phi(h) = hg$ with inverse $\phi^{-1}(h) = hg^{-1}$. Again notice ϕ is well-defined, since if $h \in \mathcal{G}_{g \cdot s}$ then $h \cdot (g \cdot s) = s$, so that $\phi(h) \cdot s = hg \cdot s = s$. □

Finally, the generalized version of Burnside's lemma given in proposition 19. We call it a lemma instead of a proposition as named below since it will be used to prove Pólya's enumeration theorem in the next section.

Proposition 36 (Burnside)**.** *Let G be a finite group acting on a set S, and let $\sim$ be the equivalence relation on S defined in proposition 34. Let S/G be the number of distinct orbits. Then the cardinality of S/G can be computed as*

$$|S/G| = \frac{1}{|G|} \sum_{g \in G} |\operatorname{Fix}(g)|.$$

Proof. With a bit of algebraic manipulation, we have

$$\begin{aligned}
|G||S/G| &= |G| \sum_{\{\mathcal{O}_s : s \in S\}} 1 \\
&= |G| \sum_{\{\mathcal{O}_s : s \in S\}} \sum_{s \in \mathcal{O}_s} \frac{1}{|\mathcal{O}_s|} \\
&= |G| \sum_{\{\mathcal{O}_s : s \in S\}} \sum_{s \in \mathcal{O}_s} \frac{|\mathcal{G}_s|}{|G|} \qquad \text{(by lemma 35)} \\
&= \sum_{\{\mathcal{O}_s : s \in S\}} \sum_{s \in \mathcal{O}_s} |\mathcal{G}_s| \\
&= \sum_{s \in S} |\mathcal{G}_s| \\
&= |\{(g, s) \in G \times S : g \cdot s = s\}| \\
&= \sum_{g \in G} |\operatorname{Fix}(g)|.
\end{aligned}$$

Therefore $|S/G|$ is computed as claimed. □

4. Finite rotations groups

With the definition of groups, we present another result that follows from the Orbit-Stabilier theorem (Lemma 35). In this section we will assume basic knowledge of matrix algebra. Recall that the 3×3 *special orthogonal group* SO_3 is defined to be the group of all rotations about the origin of $\mathbb{R}^3$.

Proposition 37. *The group SO_3 can also be thought of as the following group of matrices:*

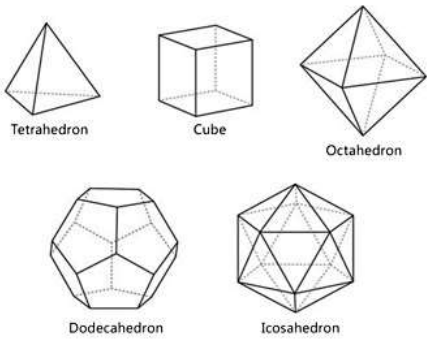
$$SO_3 := \left\{ A \in GL_3(\mathbb{R}) : AA^T = A^T A = I \text{ and } \det(A) = 1 \right\}.$$

Proof. Since elements of SO_3 are linear transformations, we can represent them by a matrix in $M \in \mathrm{GL}_3(\mathbb{R})$. Furthermore, M sends an orthonormal basis to another orthonormal basis, so it is automatic that $MM^T = M^T M = I$. Thus $\det(M) \in \{-1, 1\}$. As each such M must preserve orientation, this forces $\det(M) = 1$. □

We can think of many kinds of infinite rotation groups. The purpose of this section is to prove the following theorem.

Theorem 38. *There are only five kinds of finite subgroups of SO_3 as listed below.*

- *C_k, the cyclic group of rotations by $2\pi/k$.*
- *D_k, the dihedral group of order $2k$.*
- *T, the group of 12 rotations of a tetrahedron.*
- *O, the group of 24 rotations of a cube or octahedron.*
- *I, the group of 60 rotations of a dodecahedron or icosahedron.*



Exercise 39. *Check that the five solids above, called the Platonic solids, actually have rotation groups of order stated. In fact, the groups are A_4 and S_4 and A_5. One can also show that the Platonic solids are the only five convex regular polyhedra. Prove this using the well-known Euler's formula for polyhedra.*

Exercise 40. *Here is an easy exercise that will be used in the proof of theorem 38. Choose an axis of rotation A, and let G be a finite nontrivial group of rotations about A. Show that G is a cyclic group.*

Proof of Theorem 38. Let G be a finite nontrivial subgroup of SO_3 of order $N > 1$. Notice each nontrivial rotation in G contains exactly two poles on the unit sphere S^2, which is the intersection of the axis of rotation with S^2. Let $\mathcal{P}$ be the set of all poles of $G \setminus \{1\}$. Note that $\mathcal{P}$ can be partitioned into its orbits. To see this, we just need to see that $hp \in \mathcal{P}$ for any $h \in G \setminus \{1\}$ and $p \in \mathcal{P}$. But this is trivial: if p is stabilized by $g \in G \setminus \{1\}$, then hp is stabilized by hgh^{-1} with $hgh^{-1} \neq 1$.

For every $p \in \mathcal{P}$ let $\mathcal{G}_p$ be the group of all rotations about p that are in G, including the trivial rotation. Note that $\mathcal{G}_p$ is a subgroup of G by proposition 31, and that $\mathcal{G}_p$ is generated by a rotation of angle $2\pi/r_p$ with $r_p > 1$. Hence $|\mathcal{G}_p| = r_p$. By a simple application of proposition 34 one concludes that r_p must divide N (the assertion of this sentence is more commonly known as *Langrange's theorem*). Since there are $r_p - 1$ elements in $\mathcal{G}_p$ with exactly 2 poles, and the identity element has no poles, we have

$$\sum_{p \in \mathcal{P}} (r_p - 1) = 2(N - 1).$$

Let O_p denote the orbit of $p \in \mathcal{P}$. If there are k distinct orbits $O_1, \ldots, O_k$ in $\mathcal{P}$ with choice of representatives $p_j \in O_j$, then the above equation can be rewritten as

$$\sum_{j=1}^{k} |O_j|(r_{p_j} - 1) = 2N - 2.$$

By the Orbit-Stabilizer theorem (Lemma 35) one has

$$\sum_{j=1}^{k} \left(N - \frac{N}{r_j} \right) = 2N - 2$$

and so

$$\sum_{j=1}^{k} \left(1 - \frac{1}{r_j} \right) = 2 - \frac{2}{N}.$$

Since every term in the left hand side is at least $1/2$ and the right hand side is bounded below by 1 and bounded above by 2, we conclude that $k \in 2, 3$. It remains to investigate these two cases.

Case 1: $k = 2$. In this case

$$\frac{1}{r_1} + \frac{1}{r_2} = \frac{2}{N}.$$

As r_i divides N, one must have that $r_1 = r_2 = N$. Hence G must be the cyclic group C_N.

Case 2: $k = 3$. In this case

$$\frac{1}{r_1} + \frac{1}{r_2} + \frac{1}{r_3} = 1 + \frac{2}{N} > 1.$$

Without loss of generality let $r_3 \geq r_2 \geq r_1 > 1$. Then we must have $r_1 = 2$. As $N > 1$ we also have that $r_3 \geq 3$.

Subcase 2.1: $r_1 = r_2 = 2$ *and* $r_3 \geq 3$. Suppose $r_3 = l$. Then we will have $N = 2l$, and so $|G_{p_3}| = l$. Hence, if $\{p, p'\}$ are the poles that makes up O_3, then half of the elements of G fixes p, and the other half of G interchanges p and p'. Consequently the elements of G that fixes p must also fix p', and they are antipodal points. It is now clear that in this subcase G corresponds to the dihedral group D_l.

Subcase 2.2: $r_1 = 2$ *and* $r_3 \geq r_2 \geq 3$. We observe that we need $r_2 = 3$, and that $r_3 \leq 5$. Hence there are three possibilities.

- If $(r_1, r_2, r_3) = (2, 3, 3)$, then $N = 12$.
- If $(r_1, r_2, r_3) = (2, 3, 4)$, then $N = 24$.
- If $(r_1, r_2, r_3) = (2, 3, 5)$, then $N = 60$.

Let us just analyze the first possibilities as the others are similar. Consider r_2. The orbit corresponding to this has four points, say $\{q_1, q_2, q_3, q_4\}$. As $r_2 = 3$, up to labeling we require q_1 to be stabilized by three elements of G and q_2, q_3, q_4 to form an orbit via rotations of $2\pi/3$ around an axis through q_1. This implies that q_2, q_3, q_4 must lie on an equilateral triangle. Applying the same logic by switching out q_1 with the other q_i tells us that in fact q_1, q_2, q_3, q_4 are the vertices of the tetrahedron. Hence any rotation group of the first possibility is a subgroup of the rotation group of the tetrahedron. But the rotation group of the tetrahedron has exactly 12 elements by example 23, so this is the only possibility. □

5. Pólya's enumeration theorem

Notice Burnside's lemma tells us inequivalent symmetries (such as colorings), but does not tell us how they are counted. Fortunately there is another powerful result called Pólya's enumeration theorem. It gives us more data on our inequivalent symmetries, but is harder to compute. We first need a proposition.

Proposition 41. *Let G be a subgroup of the symmetric group $\mathfrak{S}_n$. Suppose G acts on a finite set $S = \{s_1, \ldots, s_n\}$ with n elements with the action of example 18. Define the set $X^S := \{$functions $f : S \longrightarrow X\}$ for some set $X = \{c_1, c_2, \ldots\}$. For any $f \in X^S$, also define the weight of f by*

$$x^f := \prod_{i \geq 1} x_i^{|f^{-1}(c_i)|}.$$

(a) *There is a natural action of G on X^S by following: if $g \in G$, $f \in X^S$, and $s \in S$, then $g \cdot f$ has $(g \cdot f)(s) = f(g \cdot s)$.*
(b) *If f and h are in the same orbit, then $x^f = x^h$.*
(c) *For a weak composition $\alpha = (\alpha_1, \alpha_2, \dots)$ of n, i.e. $\alpha_1 \geq \alpha_2 \geq \cdots \geq 0$ and $\sum_i \alpha_i = n$, define*

$$C_\alpha := \{f \in X^S : |f^{-1}(c_i)| = \alpha_i \text{ for all } i\}.$$

Then C_α is stable under the action of G on X^S, that is, $g \cdot f \in C_\alpha$ for every $g \in G$ and $f \in C_\alpha$. In particular, every orbit $\mathcal{O} \in X^S/G$ has $\mathcal{O} \in C_\alpha$ for some C_α.

Proof. Parts (a) and (c) are left as exercises. For part (b), write $f = g \cdot h$ for some $g \in G$. Observe that $\mathcal{G}_f$ is in bijection with $\mathcal{G}_h$ by lemma 35. Hence

$$\begin{aligned} |f^{-1}(c_i)| &= |\{s \in S : f(s) = c_i\}| \\ &= |\{g^{-1} \cdot s \in S : f(s) = c_i\}| \\ &= |\{s \in S : f(g \cdot s) = c_i = (g \cdot f)(s)\}| \\ &= |h^{-1}(c_i)|. \end{aligned}$$

A warning is that $f^{-1}(c_i)$ need not equal $h^{-1}(c_i)$ as sets. □

Definition 42. Preserve the notations in proposition 41. Define the *cycle index* of G to be

$$Z_G(t_1, t_2, \dots) = \frac{1}{|G|} \sum_{g \in G} t_1^{j_1(g)} t_2^{j_2(g)} \cdots,$$

where each $j_k(g)$ is the number of k-cycles of g. Also define the *weight function*

$$F_G(x) = \sum_{\mathcal{O} \in X^S/G} x^{\mathcal{O}},$$

where $x^{\mathcal{O}}$ is the weight x^f for any $f \in \mathcal{O}$.

Note we will manipulate polynomial series such as the ones above *formally*, and call them *generating functions*. That is, we will not worry about convergence issues like in analysis. Of course, if we want to do closed form expressions we can still use Taylor expansion if it works in a nonzero radius of convergence.

Theorem 43 (Pólya's Enumeration Theorem)**.** *With notations as in definition 42,*

$$F_G(x) = Z_G\left(\sum_{i=1}^n x_i, \sum_{i=1}^n x_i^2, \sum_{i=1}^n x_i^3, \dots\right).$$

Proof. For any $\alpha = (\alpha_1, \alpha_2, \dots)$ a weak composition of n, let us compute the fixed point for the action of some $g \in G$ on C_α. Denote this action by g_α For any $f \in C_\alpha$ to be in $\operatorname{Fix}(g_\alpha)$, we need the following conditions:

- all the elements in the orbit of any cycle of w must have the same image under f,
- the color c_i must appear α_i times as $f \in C_\alpha$.

Hence

$$|\operatorname{Fix}(g_\alpha)| = \text{the coefficient of } x^\alpha \text{ in } \prod_{k=1}^n \left(\sum_{i \geq 1} x_i^k\right)^{j_k(g)},$$

so

$$\sum_{\alpha \text{ weak composition}} |\operatorname{Fix}(g_\alpha)| x^\alpha = \prod_{k=1}^n \left(\sum_{i \geq 1} x_i^k\right)^{j_k(g)}.$$

Now, summing over all $g \in G$ and dividing by $|G|$,

$$\frac{1}{|G|} \sum_{g \in G} \sum_{\alpha} |\operatorname{Fix}(g_\alpha)| x^\alpha = \frac{1}{|G|} \sum_{g \in G} \prod_{k=1}^n \left(\sum_{i \geq 1} x_i^k\right)^{j_k(g)}.$$

The left hand side of this equation evaluates to

$$\begin{aligned}
\frac{1}{|G|}\sum_{g\in G}\sum_{\alpha}|\operatorname{Fix}(g_\alpha)|x^\alpha &= \frac{1}{|G|}\sum_{\alpha}\sum_{g\in G}|\operatorname{Fix}(g_\alpha)|x^\alpha \\
&= \sum_{\alpha}\frac{1}{|G|}\sum_{g\in G}|\operatorname{Fix}(g_\alpha)|x^\alpha \\
&= \sum_{\alpha}|C^\alpha/G|x^\alpha \qquad \text{(by proposition 36)} \\
&= \sum_{\mathcal{O}\in X^S/G} x^{\mathcal{O}} \qquad \text{(by proposition 41)} \\
&= F_G(x).
\end{aligned}$$

while the right hand side is by definition

$$\frac{1}{|G|}\sum_{g\in G}\prod_{k=1}^{n}\left(\sum_{i\geq 1}x_i^k\right)^{j_k(g)} = Z_G\left(\sum_{i=1}^{n}x_i, \sum_{i=1}^{n}x_i^2, \sum_{i=1}^{n}x_i^3, \dots\right).$$

Looking back, this is exactly what we wanted. □

Pólya's enumeration theorem tells us that $F_G(x)$ is the generating function for the number of colorings under the action of G, which can be computed just by looking at the cycle index. The advantage is that once we know what G we are dealing with, we simply substitute it into the cycle index and we are done. It is easy to get back Burnside's lemma from Pólya's enumeration theorem by substituting each $x_i = 1$.

Example 44. Let us look back at example 20. Since $G = D_4$ in this case, by Pólya's enumeration theorem, and noticing that $|X| = 3$ (so we treat $0 = x_4 = x_5 = \cdots$) and $n = 4$ in this case, the generating function for the number of colorings are

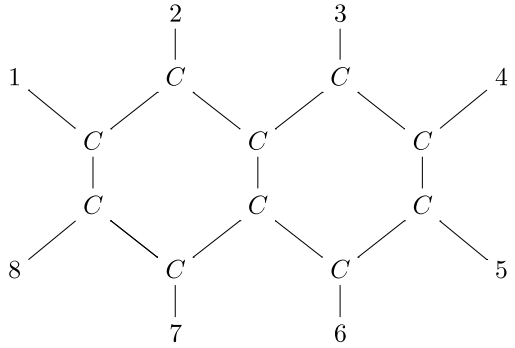
$$\begin{aligned}
\frac{1}{8}\sum_{g\in D_4}\prod_{k=1}^{4}(x_1^k + x_2^k + x_3^k)^{j_k(g)} &= \frac{1}{8}\big((x_1+x_2+x_3)^4 + 2(x_1+x_2+x_3)^2(x_1^2+x_2^2+x_3^2) \\
&\qquad + 3(x_1^2+x_2^2+x_3^2)^2 + 2(x_1^4+x_2^4+x_3^4)\big) \\
&= x_1^4 + x_1^3x_2 + x_1^3x_3 + \cdots + x_2x_3^3 + x_3^4.
\end{aligned}$$

The advantage of this generating function is that it tells how many many of each coloring there are. For example the x_1^4 and $x_2x_3^3$ term tells us there is only one coloring of the square with all vertices colored the first color, or one vertices colored the second and the rest colored with color three. To get back Burnside's lemma, simply substitute $x_1 = x_2 = x_3 = 1$, and we still get the correct count of 21 colorings.

Let us look at some more examples of Pólya's enumeration theorem. Three will be outlined with some details left to the reader, and the remaining two given as exercise.

Example 45. Let us look back at example 25. Say we are interested in the more general question of finding the number of isomers of the following molecule, where the numbers are to be filled k number of H and $8-k$

number of CH_3 for some fixed $k \in \{0, \ldots, 8\}$.



The symmetry group is still the subgroup

$$V := \{1, (1\ 5)(2\ 6)(3\ 7)(4\ 8), (1\ 4)(2\ 7)(3\ 6)(5\ 8), (1\ 8)(2\ 3)(4\ 5)(6\ 7)\}$$

of $\mathfrak{S}_8$, and the cycle index of V is

$$Z_V(t_1, t_2) = \frac{1}{4}(t_1^8 + 3t_2^4).$$

By a similar reasoning as example 40, the polynomial we want to look at is

$$\begin{aligned} Z_V(1+t, 1+t^2) &= \frac{1}{4}((1+t)^8 + 3(1+t^2)^4) \\ &= 1 + 2x + 10x^2 + 14x^3 + 22x^4 + 14x^5 + 10x^6 + 2x^7 + x^8. \end{aligned}$$

From here we can read off what we want using the coefficients. For example, the number of isomers of tetramethylnaphthalene is the coefficient of x^4, which is 22, agreeing with our computation in example 25. As another example, there are two isomers of naphthol, which is obtained by filling in the numbers in the molecule above with seven H one hydroxyl group (OH).

Exercise 46 (Enumerating isomers of alkane)**.** *An alkane is a molecule with n number of carbon (C) atoms and $2n + 2$ number of hydrogen (H) atoms such that every carbon atom is joined to exactly four hydrogen atoms, and every hydrogen atom is joined to exactly one carbon atom. A substituted alkane is an alkane where a hydrogen atom is replaced by some X.*

(a) Let a_n be the number of isomers of substituted alkanes with exactly n carbon atoms, and consider the generating function

$$A(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \cdots.$$

Show that

$$A(x) = \frac{x}{6}\left((A(x))^3 + 3A(x)A(x^2) + 2A(x^3)\right),$$

and explain how we can (possibly recursively) compute the values a_n using the above relation.

(b) Let b_n be the number of isomers of alkanes with exactly n carbon atoms. By considering the cycle index of $\mathfrak{S}_3$ or otherwise, set up a recursion to compute the values b_n. It may be useful to consider the generating function $B(x) = b_0 + b_1 x + b_2 x^2 + b_3 x^3 + \cdots$.

Example 47. A classical application of Pólya's enumeration theorem is to count the number of necklaces composed of n beads of two colors, and furthermore find out how many of these necklaces have k beads of the first color. In this case $G = \mathfrak{R}_n$ since two necklaces are equivalent if they are the same under rotation (not reflection!). We need to compute $Z_G(t_1, \ldots, t_n)$ first. It is easy to see that, again letting $r := (1\ \cdots\ n) \in \mathfrak{R}_n$, there are only $\gcd(k, n)$ cycles of length $n/\gcd(k, n)$ in r^k for $1 \le k \le n$. Hence

$$Z_G(t_1, \ldots, t_n) = \frac{1}{|G|}\sum_{k=1}^{n} t_{n/\gcd(k,n)}^{\gcd(k,n)} = \frac{1}{n}\sum_{d|n} \varphi(n/d) t_{n/d}^d = \frac{1}{n}\sum_{d|n} \varphi(d) t_d^{n/d},$$

where $\varphi(d) := \{m \in \{1, \ldots, d\} : \gcd(m, d) = 1\}$ is called *Euler's totient function.* Here we implicitly use the easy fact that $\gcd(k, n) = g$ if and only if $\gcd(k/g, n/g) = 1$.

Since we have two colors, call a color x and the other 1. We will not need a new variable for another color following the argument completely. Anyway, by Pólya's enumeration theorem, the polynomial we want to look at is

$$Z_G(1+t, 1+t^2, \ldots, 1+t^n) = \frac{1}{n}\sum_{d|n} \varphi(d)(1+t^d)^{n/d}$$

This already tell us, by substituting $t = 1$, that the number of necklaces with n beads of two colors is

$$Z_G(2, \ldots, 2) = \frac{1}{n}\sum_{d|n} \varphi(d)2^{n/d}$$

To find those necklaces that has k beads of the first color, we need to look at the coefficient of t^k in $Z_G(1+t, 1+t^2, \ldots, 1+t^n)$, which is

$$\frac{1}{n}\sum_{d|\gcd(k,n)} \varphi(d)\binom{n/d}{k/d}.$$

This is because it suffices to look at those d that divides k, and since d divides n it must divides $\gcd(k, n)$. The binomial coefficient is combinatorial observation.

Exercise 48. *A graph is a pair of finite sets (V, E) such that $E \subset V \times V$. Two graphs (V_1, E_1) and (V_2, E_2) are isomorphic if there exists a bijection $\varphi : V_1 \longrightarrow V_2$ such that $(v, w) \in E_1$ if and only if $(\varphi(v), \varphi(w)) \in E_2$. Classify graphs with at most 5 vertices up to isomorphism, i.e. in the classification no two graphs should be isomorphic. Include the empty graph $(\emptyset, \emptyset)$ as well. Do it for more than 5 vertices for more computational practice.*

Example 49. We can use Pólya's enumeration theorem to give a beautiful identity (which is the third equation below). For variables $r_1, \cdots, r_n$, write $p_k = \sum_{i=1}^n r_i^k$. Then

$$\prod_{i=1}^n \frac{1}{1-r_i t} = \sum_{m\geq 0} F_{\mathfrak{S}_m}(r_1, \ldots, r_n)t^m = \sum_{m\geq 0} Z_{\mathfrak{S}_m}(p_1, \ldots, p_n)t^m.$$

The left equality is true since $1/(1-r_i t) = 1 + r_i t + (r_i t)^2 + (r_i t)^3 + \cdots$, and the right equality is Pólya's enumeration theorem. Also noting that

$$\frac{1}{1-r_i t} = \exp\left(\ln\frac{1}{1-r_i t}\right) = \exp\left(r_i t + \frac{r_i^2 t^2}{2} + \frac{r_i^3 t^3}{3} + \cdots\right),$$

by summing over i we have

$$\sum_{m\geq 0} Z_{\mathfrak{S}_m}(p_1, \ldots, p_n)t^m = \exp\left(p_1 t + \frac{p_2 t^2}{2} + \frac{p_3 t^3}{3} + \cdots\right).$$

Notice this means that the right hand side gives us the cycle index polynomials of any symmetric group $\mathfrak{S}_m$.

Exercise 50. *Let $k \in \mathbb{Z}_{>0}$. Prove that the number $f_k(n)$ of permutations $\sigma \in \mathfrak{S}_n$, all of whose cycle lengths are divisible by k in its cycle factorization, is given by*

$$f_k(n) = \begin{cases} (n-1)! \displaystyle\prod_{j=1}^{n/k-1} \frac{jk+1}{jk} & \text{if } k|n \\ 0 & \text{otherwise.} \end{cases}$$

For more examples of applications of Pólya theory, one may refer to the beautifully written book [5] written by George Pólya himself together with another collaborator of his.

6. Cayley's theorem

We have summarized Pólya theory. There is still one question that is theoretically important, but practically not so. We only did Pólya's enumeration theorem over permutation groups. Why not do it for others? It turns out this is sufficient by the following theorem. For a set G, let us define $\mathfrak{S}_G$ to be the set of all bijective functions from G to G. In particular, a finite group G is isomorphic to a subgroup of $\mathfrak{S}_{|G|}$. Just like the symmetric group over a finite set, $\mathfrak{S}_G$ is a group under function composition as well.

Theorem 51 (Cayley)**.** *Every group G is isomorphic to a subgroup of $\mathfrak{S}_G$, i.e. we can construct an injection $f : G \longrightarrow \mathfrak{S}_G$ such that $f(g * h) = f(g)f(h)$ for every $g, h \in G$.*

Proof. Observe $\mathfrak{S}_G$ has a subgroup $\lambda := \{\lambda_g\}_{g\in G}$, where each $\lambda_g : G \longrightarrow G$ is the *left multiplication map* $\lambda_g(h) = gh$. Each $\lambda_g : G \longrightarrow G$ is a bijection with inverse given by $\lambda_g^{-1}(h) = g^{-1}h$, and it is easy to see λ is a subgroup of $\mathfrak{S}_G$.

Let us define the map $\varphi : G \longrightarrow \mathfrak{S}_G$ by $\varphi(g) = \lambda_g$, which satisfies the property that $\varphi(g * h) = \varphi(g)\varphi(h)$ for every $g, h \in G$. It remains to check φ is injective. If $\varphi(g) = \varphi(h)$ for some $g, h \in G$, then $\lambda_g \equiv \lambda_h$. In particular, $g = \lambda_g(1) = \lambda_h(1) = h$. □

7. The Futurama Theorem

In this section we will apply the idea of the symmetric group to solve a problem which is the underlying theme of the acclaimed Futurama episode "The Prisoner of Benda". This theorem was thought up by Ken Keeler, a writer for Futurama who holds a Ph.D in applied mathematics. He directed the Futurama episode named above, and according to him a main aim of the episode was to popularize math among young people.

We will now see the Futurama theorem requires no more than manipulation of elements of the symmetric group. In order to not spoil the episode in case you haven't watch it, I've rephrased the theorem as below.

Theorem 52 (Futurama)**.** *Suppose we have a group of n people, with $n \in \mathbb{Z}_{\geq 0}$. Each person has exactly one gift, and exchanges gifts with one another pairwise for as many times as they wish. Suppose in addition that every two person can switch gifts with each other at most once (even if they had different gifts before). If we add in two extra person, each with a gift, then it is possible to pairwise exchange gifts among these $n + 2$ people under the added assumption such that all $n + 2$ people get back their own gifts.*

Proof. The case for $n = 0$ is trivial. Suppose $n \geq 1$. Let $\{1, \ldots, n + 2\}$ be the set of $n + 2$ people, and $\{1, \ldots, n\}$ be the subset of n people. We translate the statement of the theorem as follow: given $\sigma \in \mathfrak{S}_n$, consider it as an element in $\mathfrak{S}_{n+2}$ with $n + 1$ and $n + 2$ fixed. Then there is a set of pairwise distinct transpositions $\{\tau_j\}_{j=1}^k$ of $\mathfrak{S}_{n+2}$ such that each $\tau \in \{\tau_j\}_{j=1}^k$ does not fix both of $n + 1$ and $n + 2$, and $\tau_k\tau_{k-1}\cdots\tau_2\tau_1\sigma$ is the identity in $\mathfrak{S}_{n+2}$.

Let $\sigma = \sigma_1 \cdots \sigma_l$ be the cycle factorization of $\sigma \in \mathfrak{S}_n$. Consider one of the σ_i, with $1 \leq i \leq l$. Without loss of generality, suppose $\sigma_i = (1\ \ 2\ \ \cdots\ \ m)$, where $1 \leq m \leq n$. This is valid since if σ_i is a cycle with different index, reindex the elements in its cycle into the form above. Consider σ_i to be an element in $\mathfrak{S}_{n+2}$ with $n+1$ and $n + 2$ fixed. Now consider the set of transpositions $\{(n + 1\ \ q) : 1 \leq q \leq m\} \sqcup \{(n + 2\ \ m), (n + 2\ \ 1)\}$. Let

$$\kappa_i := (n + 1\ \ 1)(n + 1\ \ 2)\cdots(n + 1\ \ m - 1)(n + 2\ \ m)(n + 1\ \ m)(n + 2\ \ 1).$$

Then we can easily check that $\kappa_i\sigma_i = (n + 1\ \ n + 2)$ in $\mathfrak{S}_{n+2}$. Now letting $\kappa := \kappa_l\kappa_{l-1}\cdots\kappa_2\kappa_1$, we see inductively that $\kappa\sigma = (n+1\ \ n+2)^l$ in $\mathfrak{S}_{n+2}$, as $(n+1\ \ n+2)$ commutes with each $\sigma_1, \ldots, \sigma_l$ since they fix $n+1$ and $n+2$. Thus if l is even, multiply κ to σ to get $\kappa\sigma = e$ in $\mathfrak{S}_{n+2}$, and if l is odd, multiply the chain of transpositions $(n + 1\ \ n + 2)\kappa$ to σ instead.

To make sure our definition of κ works, we need to check each of its transpositions are pairwise distinct. But each κ_i has pairwise distinct transpositions by definition, $(n + 1\ \ n + 2)$ does not appear in κ, and transpositions from two different κ_i and κ_j are pairwise distinct too since the corresponding disjoint cycles σ_i and σ_j permutes pairwise distinct elements. □

Note that in the proof of the Futurama theorem, we exchange gifts by "fixing" the gifts and exchange the two person around instead. This helps us keep track of pairs of people that has already exchanged gifts with each other so we do not violate the given condition. Also, by showing that introducing two person is sufficient, this implies introducing more than two person also works.

8. The 100 Prisoners Problem

The 100 Prisoners Problem is the following problem.

The 100 Prisoners Problem. A crazy warden has thought up a game for 100 prisoners he has. He goes into an empty room and makes up 100 identical boxes in a row, inside writing each of the 100 prisoners' (unique) names. Each prisoner is, once at a time, allowed to enter the room and open 50 boxes. After each prisoner's turn, the boxes are all closed and the next prisoner comes in, until the last one finishes his turn. If every prisoner opens the box that has his/her name in it, all are free to go. Else all prisoners gets killed. The rule of the game is that the 100 prisoners can only have a discussion before the game starts, and no more communication of any kind is allowed after the game starts.

We aim to find a good strategy for the problem above. Note that choosing the boxes randomly for each prisoner is not a very good strategy, since each box has a 2^{-1} chance of being opened and we open 50 boxes out of 100, so the probability of survival for each prisoner is

$$(2^{-1})^{50}(1-2^{-1})^{50} = 2^{-100} \approx 7.888 \cdot 10^{-31},$$

which is exactly 1 in 2^{100}, an extremely bad chance. There is a strategy that has a survival probability that is at least 10^{29} times better, which we state its probability of survival now.

Proposition 53. *There exists a strategy for the 100 Prisoners Problem with probability of survival*

$$1 - \sum_{j=51}^{100} \frac{1}{j} \approx 0.312,$$

which is about 1 *in* 3.207*.*

Before proving this proposition by giving a strategy for it, we need a lemma.

Lemma 54. *Let k and n be positive integers with $k \leq n$. The number of k-cycles in $\mathfrak{S}_n$ is*

$$\binom{n}{k}(k-1)!.$$

Proof. Easy combinatorics exercise. □

It remains to apply lemma 54 in a clever way to prove proposition 53.

Proof of Proposition 53. We describe the strategy first before proving its probability of survival is as claimed. Without loss of generality, name the prisoners integers 1 to 100, and have the prisoners memorize everyone else's names. Also have the prisoners view the boxes in the room as integers 1 to 100 lined up in a row. For each $k \in \{1, \ldots, 100\}$, when prisoner k goes into the room, he/she will open box k first. If box k has his/her name in it, the prisoner end the turn. Else the prisoner will see the name of prisoner n_1 in box k, and the prisoner will open box n_1. The prisoner repeats this procedure until he/she opens a box with his/her name in it or the 50^{th} box, whichever is earlier.

Next we calculate the probability of survival. Note that the above strategy is equivalent to counting the number of elements in the symmetric group $\mathfrak{S}_{100}$ that has each cycle of length at most 50 in its cycle factorization, since if $\sigma \in \mathfrak{S}_{100}$, then prisoner k chooses the boxes to open in the fashion $k, \sigma(k), \sigma^2(k), \ldots$, so prisoner k does not open the box containing his/her name if and only if $k \notin \{\sigma^l(k) : l \in \mathbb{Z}_{>0}\}$. However, it is not easy to count the number of elements in $\mathfrak{S}_{100}$ that has each cycle of length at most 50 in its cycle factorization directly, so we instead count the number of elements $\pi \in \mathfrak{S}_{100}$ that has a cycle of length at least 51 in its cycle factorization, i.e. the probability of death. Note that we can only have one such cycle in σ, else if not then σ permutes at least 102 elements, contradiction. Let the cycle in σ of length at least 51 be r. There are $\binom{100}{r}(r-1)!$ ways to choose this cycle by lemma 54, and $(100-r)!$ ways to permute the other $100-r$ elements of $\{1, \ldots, 100\}$, so there are

$$\binom{100}{r}(r-1)!(100-r)! = \frac{100!(r-1)!(100-r)!}{r!(100-r)!} = \frac{100!}{r}$$

permutations $\sigma \in \mathfrak{S}_{100}$ that has a cycle of length $r \in \{51, \dots, 100\}$ in its cycle factorization. Summing over r, the probability of death is thus

$$\frac{1}{|\,\mathfrak{S}_{100}\,|} \sum_{r=51}^{100} \binom{100}{r} (r-1)!(100-r)! = \frac{1}{100!} \sum_{r=51}^{100} \frac{100!}{r} = \sum_{r=51}^{100} \frac{1}{r}.$$

Hence the probability of survival is 1 minus the value above, as claimed. □

Proposition 53 can be generalized in an obvious way. Proving this requires some calculus.

Lemma 55. *We have*

$$0.3 < 1 - \log 2,$$

where note that the logarithmic function above is in base e.

Proof. Easy exercise. □

Lemma 56. *There exists a constant $\gamma \in \mathbb{R}$, called Euler's constant, such that*

$$\lim_{n\to\infty} \left(-\log n + \sum_{j=1}^{n} \frac{1}{j} \right) = \gamma,$$

where note that the logarithmic function above is in base e. □

Proof. A calculus exercise. □

Theorem 57. *Change the 100 Prisoners Problem to the case of $2n$ prisoners and n boxes for some positive integer n. Then the probability of survival for each prisoner is*

$$1 - \sum_{j=n+1}^{2n} \frac{1}{j},$$

and furthermore this value is always greater than 0.3.

Proof. The first part has an exact same proof as proposition 53. For the second part, note that $\sum_{j=n+1}^{2n} \frac{1}{j}$ is monotone increasing if viewed as a function in n, so $1 - \sum_{j=n+1}^{2n} \frac{1}{j}$ is monotone decreasing if viewed as a function in n. Also note that for any positive integer n,

$$0.5 = 1 - \frac{1}{2} = 1 - \sum_{j=1+1}^{2\cdot 1} \frac{1}{j} \geq 1 - \sum_{j=n+1}^{2n} \frac{1}{j} \geq 1 - \sum_{j=n+1}^{2n} \frac{1}{n+1} \geq 1 - \frac{2n-n}{n+1} \geq 0,$$

so that the sequence $\{1 - \sum_{j=n+1}^{2n} \frac{1}{j}\}_{n\in\mathbb{Z}_{>0}}$ is bounded, and hence converges by the Monotone Convergence Theorem. Thus it follows by sum of finitely many convergent limits, basic properties of logarithms, and monotonicity of $\{1 - \sum_{j=n+1}^{2n} \frac{1}{j}\}_{n\in\mathbb{Z}_{>0}}$ that

$$\begin{aligned}
1 - \sum_{j=n+1}^{2n} \frac{1}{j} &\geq \lim_{n\to\infty} \left(1 - \sum_{j=n+1}^{2n} \frac{1}{j} \right) \\
&= \gamma - \gamma + \lim_{n\to\infty} \left(1 - \sum_{j=1}^{2n} \frac{1}{j} + \sum_{j=1}^{n} \frac{1}{j} \right) \\
&= \lim_{n\to\infty} \left(-\log 2n + \sum_{j=1}^{2n} \frac{1}{j} \right) - \lim_{n\to\infty} \left(-\log n + \sum_{j=1}^{n} \frac{1}{j} \right) + \lim_{n\to\infty} \left(1 - \sum_{j=1}^{2n} \frac{1}{j} + \sum_{j=1}^{n} \frac{1}{j} \right)
\end{aligned}$$

$$
\begin{aligned}
&= \lim_{n\to\infty} \left(\left(-\log 2n + \sum_{j=1}^{2n} \frac{1}{j} \right) - \left(-\log n + \sum_{j=1}^{n} \frac{1}{j} \right) + \left(1 - \sum_{j=1}^{2n} \frac{1}{j} + \sum_{j=1}^{n} \frac{1}{j} \right) \right) \\
&= \lim_{n\to\infty} \left(1 - (\log 2n - \log n)\right) \\
&= \lim_{n\to\infty} \left(1 - \log \frac{2n}{n}\right) \\
&= 1 - \log 2. \\
&> 0.3,
\end{aligned}
$$

where γ is Euler's constant as in lemma 56, and the last inequality is by lemma 55. □

Using a calculator, $1 - \log 2 \approx 0.307$, so the lower bound of 0.3 is a very good bound for theorem 57 that we can prove just by elementary single-variable calculus.

You may wonder if the strategy given in the proof of proposition 53 is *optimal*, i.e. the probability of survival for each prisoner is maximal, as that is probably the most important in the prisoners' mind. Unfortunately this is true, so the probability of each prisoner surviving in the 100 Prisoners Problem is not very high. The following exercise explains why.

Exercise 58. *Consider an alternative game, still with the* $2n$ *prisoners and* $2n$ *boxes labeled integers* 1 *to* $2n$*, as follows. Let prisoner* 1 *go into the room and open boxes until he/she opens the box with his/her name. If all boxes are opened after that, the game ends. Else the warden tells the prisoner with the lowest number among all the boxes not opened to come in and open boxes until he/she finds the boxes with his/her name (note the boxes opened by prisoner* 1 *is still left opened). Continue the procedure until all boxes are opened. The prisoners all survive if no prisoner opened more than* n *boxes, else all the prisoners die.*

(a) Show that the alternative game terminates, i.e. all $2n$ *boxes will be opened after the* k^{th} *prisoner for some* $1 \leq k \leq 2n$*.*

(b) Show that the alternative game is strategy-neutral, i.e. the probability of survival for each prisoner is the same no matter what strategy the prisoner use to open the box.

(c) We can change the 100 prisoners problem by letting them open the box until each prisoner finds the box with his/her name in it, and that the prisoners win iff no prisoner opened more than 50 boxes. Clearly this does not change the chance of survival for the prisoners. Show that the alternative game dominates the modified 100 Prisoner Problem, i.e. every strategy in the 100 Prisoners Problem can be directly implemented in the alternative game, and furthermore every winning strategy in the 100 Prisoners Problem is a winning strategy in the alternative game.

(d) Compute the probability of survival for each prisoner in this alternative game.

(e) Conclude that the strategy given in the proof for the 100 Prisoners Problem is optimal.

References

[1] Mark Armstrong, *Groups and Symmetry*, Springer-Verlag New York, 1988.

[2] Michael Artin, *Algebra (Second Edition)*, Pearson, 2010.

[3] Eugene Curtin and Max Warshauer, *The Locker Puzzle*, Mathematical Intelligencer **28** (28–31), 2006.

[4] David Dummit and Richard Foote, *Abstract Algebra (Third Edition)*, John Wiley and Sons, 2004.

[5] George Pólya and Ronald Read, *Combinatorial enumeration of groups, graphs, and chemical compounds*, Springer-Verlag, 1987.

[6] Richard Stanley, *Enumerative Combinatorics Volume 2*, Cambridge Studies in Advanced Mathematics, 2001.