Permutations as Products of Transpositions

George Mackiw

When writing a permutation as a product of transpositions, what is the smallest number of transpositions that can be used? This question and variants of it occur both abstractly [2] and in applied settings such as data exchange and sorting [3]. The answer is known and easily stated: the minimum number is precisely $n - r$, where $r$ is the number of disjoint cycles in the given permutation on $n$ letters. One way to establish this result is to use an inductive argument relying on an analysis of how cycles multiply [1]. Another approach, employed in [4], restates the problem in the language of graph theory and makes use of the fact that a connected graph with $n$ vertices must have at least $n - 1$ edges.

The purpose of this note is to provide an alternate derivation of this result that uses only elementary linear algebra. The very answer, $n - r$, seems to suggest some dimension counting involving complementary spaces, and, indeed, our argument takes advantage of orthogonality and the Gram-Schmidt process.

The linear algebraic connection is a natural one. Elements of the symmetric group $S_n$ permute coordinates in $\mathbb{R}^n$ and are often realized as permutation matrices. More precisely, we regard a permutation $\sigma$ in $S_n$ as acting on the Euclidean space $\mathbb{R}^n$ by $\sigma e_i = e_{\sigma(i)}$, where $e_1, e_2, \ldots, e_n$ denotes the natural basis of $\mathbb{R}^n$.

In this setting, transpositions have a simple geometric interpretation. Given the transposition $\tau = (i, j)$ in $S_n$, $i < j$, we call the vector $e_i - e_j$ in $\mathbb{R}^n$ the vector associated to $\tau$. Notice that $\tau$ sends this vector to its negative $e_j - e_i$. Further, $\tau$ fixes pointwise the collection of $n - 1$ vectors $\{e_k | k \neq i, j\} \cup \{e_i + e_j\}$ which are all orthogonal to $e_i - e_j$. Indeed, these $n - 1$ vectors form a basis for the subspace (hyperplane) orthogonal to the vector $e_i - e_j$. Simply put, $\tau$ acts as the reflection through the hyperplane orthogonal to $e_i - e_j$.

We also attach a subspace to any permutation $\sigma$—the fixed point space, $V_\sigma$, consisting of all vectors $x$ in $\mathbb{R}^n$ with $\sigma x = x$. This is the eigenspace of $\sigma$ corresponding to the eigenvalue $\lambda = 1$. It always has positive dimension since, for example, any vector all of whose components are equal is in $V_\sigma$ for any $\sigma$. As a matter of fact, it is not difficult to see that the fixed point space is determined by the cycle structure of the permutation. Note that vectors in $V_\sigma$ must have their $i$th and $j$th components agreeing whenever $i$ and $j$ occur in a common cycle of $\sigma$. It follows that if $\sigma$ is written as a product of $r$ disjoint cycles, including trivial cycles containing only one point, then $V_\sigma$ is $r$-dimensional with each cycle of $\sigma$ contributing a basis element in a natural way to $V_\sigma$.

For example, the permutation $\sigma = (2, 5, 3)(1, 6)$ in $S_7$ has a four-dimensional fixed point space in $\mathbb{R}^7$ that has the vectors $e_2 + e_5 + e_3, e_1 + e_6, e_4$ and $e_7$ for a basis.
We read products of permutations from right to left. Thus, \((1, 2, 3, 4, 5) = (1, 5)(1, 4)(1, 3)(1, 2)\) expresses a five-cycle as a product of four transpositions. In like fashion, an \(s\)-cycle can be written using \(s - 1\) transpositions. So, a permutation in \(S_n\) consisting of \(r\) cycles can be written as a product of \(n - r\) transpositions. We are now ready to show that no fewer number of transpositions can be employed. Note that in counting cycles of a permutation we always include trivial one element cycles.

**Theorem 1.** A permutation in \(S_n\) cannot be written as the product of fewer than \(n - r\) transpositions, where \(r\) is the number of disjoint cycles in the permutation.

**Proof:** Suppose \(\sigma\) in \(S_n\) is written as \(\sigma = \tau_1 \tau_2 \cdots \tau_k\), where the \(\tau_i\)'s are transpositions. Viewing transpositions as reflections through hyperplanes, let \(v_i, i = 1, 2, \ldots, k\), be the vectors associated to these transpositions. Recall that \(v_i\) is a vector orthogonal to the hyperplane determined by \(\tau_i\). The Gram-Schmidt orthogonalization process guarantees the existence of at least \(n - k\) linearly independent vectors that are orthogonal to the subspace spanned by the \(v_i\)'s. These \(n - k\) vectors thus lie in the intersection of the \(k\) hyperplanes determined by the transpositions and are thus pointwise fixed by each of the transpositions. Thus these vectors are fixed by \(\sigma\), and so \(\dim V_\sigma \geq n - k\). But, \(\dim V_\sigma = r = \) number of cycles in \(\sigma\). The result \(k \geq n - r\) follows. \(\square\)

Whenever \(\sigma = \tau_1 \tau_2 \cdots \tau_k\), a product of transpositions, and \(k\) is the minimum number allowed by Theorem 1, we refer to this as a minimal representation of \(\sigma\). We now use orthogonality to show that a minimal representation must have associated vectors that are linearly independent.

Any \(\sigma\) in \(S_n\) determines a direct sum decomposition \(R^n = V_\sigma \oplus V_\sigma\perp\), where \(V_\sigma\perp\) denotes the orthogonal complement in \(R^n\) of the fixed point space \(V_\sigma\). If \(\sigma = \tau_1 \tau_2 \cdots \tau_k\) is a minimal representation, then \(\dim V_\sigma = n - k\). Now the vectors \(v_i, i = 1, 2, \ldots, k\), associated to the transpositions \(\tau_i\) are normal vectors to hyperplanes \(H_i\). Since \(\tau_i\) fixes \(H_i\) pointwise, the intersection \(\bigcap_{i=1}^k H_i\) is a subspace contained in \(V_\sigma\). The intersection of these \(k\) hyperplanes is the solution space to a \(k\) by \(n\) homogeneous system of equations, where the \(i\)th equation expresses the requirement that a vector in \(H_i\) must be orthogonal to \(v_i\). Elementary results concerning rank and solution spaces of systems of equations show that \(\dim(\bigcap_{i=1}^k H_i) \geq n - k\), with equality occurring exactly when the normal vectors \(v_1, v_2, \ldots, v_k\) are linearly independent. We have derived the following result.

**Theorem 2.** If the representation \(\sigma = \tau_1 \tau_2 \cdots \tau_k\) is a minimal one, then the associated vectors \(v_1, v_2, \ldots, v_k\) are linearly independent and form a basis for \(V_\sigma\perp\). \(\square\)

For example, \((1, 6)(3, 4)(4, 6)(1, 3)\) could not be a minimal representation, due to the dependence relation \(e_1 - e_6 = (e_3 - e_4) + (e_4 - e_6) + (e_1 - e_3)\).

Other reasonably intuitive results about minimal products of transpositions can be obtained using this approach. For example, a minimal representation \(\sigma = \tau_1 \tau_2 \cdots \tau_k\) must respect the cycle structure of \(\sigma\). For, suppose that some transposition \(\tau_i = (a, b)\) was such that \(a\) and \(b\) belonged to different cycles of \(\sigma\). Then the vector \(v = e_a - e_b\) would not have inner product zero with the vector \(w = \sum e_\alpha\), where \(\alpha\) ranges through the elements of the cycle of \(\sigma\) containing \(a\). But this contradicts our result that \(w\) is in \(V_\sigma\), while the associated vector \(v\) is in \(V_\sigma\perp\).
particular, no transposition \( \tau_i = (a, b) \) can have either \( a \) or \( b \) belonging to a trivial one element cycle of \( \sigma \).

The converse of Theorem 2 is also true, though we omit the arguments.

REFERENCES


Dept. of Mathematical Sciences
Loyola College in Maryland
Baltimore, MD 21210
mackiw@loyola.edu

---

**Congruences Relating the Order of a Group to the Number of Conjugacy Classes**

**Bjorn Poonen**

Let \( G \) be a finite group, and let \( |G| \) denote its order. Let \( s \) be the number of conjugacy classes in \( G \). Burnside, in his 1911 text on the theory of finite groups, used representation theory to prove that if \( |G| \) is odd, then \( |G| \equiv s \) (mod 16). (See p. 295 of [1].) On p. 320 of the same book, he left as an exercise to show that if every prime dividing \( |G| \) is congruent to 1 modulo 4, then \( |G| \equiv s \) (mod 32). The purpose of this note is to show how elementary counting arguments can yield other congruences in the same spirit. Here is what we will prove:

**Theorem.** Let \( m \geq 2 \) be an integer. If each prime divisor of \( |G| \) is congruent to 1 modulo \( m \), then \( |G| \equiv s \) (mod \( 2m^2 \)).

Taking \( m = 2 \) in this theorem yields only \( |G| \equiv s \) (mod 8), which is weaker than Burnside's original result. On the other hand, taking \( m = 4 \) yields exactly his exercise.

**Proof:** Let

\[
T = \{(g, h) \in G \times G | gh \neq hg\}.
\]

For each unordered pair \( \{C_1, C_2\} \) of cyclic subgroups of \( G \), we may consider the set of \( (g, h) \) in \( G \times G \) such that the subgroups \( \langle g \rangle, \langle h \rangle \) they generate are \( C_1 \) and \( C_2 \) in some order. Such subsets clearly form a partition of \( G \times G \).