Derangements in Sylow Subgroups of Symmetric Groups

Brian Peterson and Linda Valdés
Department of Mathematics
San José State University
San José, CA
USA 95192

ABSTRACT. Let $H_n < S_n$, where H_n is a Sylow p-subgroup of S_n, the symmetric group on n letters. Let h_n denote the number of derangements in H_n, and $f_n = \frac{h_n}{|H_n|}$. We will show that the sequence $\{f_n\}_{n=1}^{\infty}$ is dense in the unit interval, but is Cesàro convergent to 0.

1 Introduction

Let S_n denote the symmetric group on n letters, i.e. the group of all permutations of the set $X_n = \{1, \ldots, n\}$ for $n \geq 1$. Given a subgroup $H_n < S_n$ let h_n denote the number of derangements in H_n which act on X_n with no fixed points. Roger Alperin [1] has asked whether there are interesting families of subgroups $H_n < S_n$ (with $H_n \subseteq H_{n+1}$ when we view S_n as a subgroup of S_{n+1} in the obvious way) for which $\lim_{n\to\infty} \frac{h_n}{|H_n|}$ exists. He uses exponential generating functions to show that this limit equals $\frac{1}{e}$ if each $H_n = S_n$ (well known), and again if each $H_n = A_n$, the alternating group. We will let p be an arbitrary prime which is fixed throughout and take H_n to be a Sylow p-subgroup of S_n, and show that this limit does not exist. Indeed, after introducing a probabilistic point of view, we will find a formula for $f_n = \frac{h_n}{|H_n|}$ and show that the sequence $\{f_n\}_{n=1}^{\infty}$ is dense in the unit interval, but is Cesàro convergent to 0. Note that the value of f_n does not depend on the choice of H_n since all Sylow p-subgroups of S_n are conjugate, and conjugate elements of S_n have the same cycle structure.

We begin by establishing some notation. Let $m = \lfloor \frac{\ln n}{\ln p} \rfloor$ and let the base p expansion of n be given by $n = \sum_{i=0}^{m} d_i p^i$, where each $d_i \in \{0, \ldots, p-1\}$. Note that $m = 0$ iff $n < p$ iff $p \nmid n!$, and we may regard this case as trivial,

since then a Sylow p-subgroup of S_n is trivial with no derangements and $f_n = 0$. Most of what follows is true for $m = 0$ though often vacuous in this case; but, in any event, we may assume henceforth that $m \geq 1$, i.e. $n \geq p$. Let $e_i = \lfloor \frac{n}{p^i} \rfloor$ for $0 \leq i \leq m$ and let $e = \sum_{i=1}^{m} e_i$. It is well known that $p^e \| n!$, i.e. $p^e \mid n!$ and $p^{e+1} \nmid n!$.

2 Construction of Sylow p-subgroups

The basis of our approach is to consider a certain family $\mathbf{B}$ of subsets of X_n. For integers $i \geq 0$, if the natural numbers are partitioned into consecutive blocks of length p^i, the jth block is $B_{i,j} = \{(j-1)p^i + 1, \ldots, jp^i\}$. We will use the notation $B_{i,j} = [(j-1)p^i + 1, jp^i]$. In particular, $B_{o,j} = [j]$. Now $B_{i,j} \subseteq X_n$ iff $jp^i \in X_n$ iff $0 \leq i \leq m$ and $1 \leq j \leq e_i$. Let $\mathbf{B}_i = \{B_{i,j} \mid 1 \leq j \leq e_i\}$ for $0 \leq i \leq m$ and let $\mathbf{B} = \bigcup_{i=0}^{m} \mathbf{B}_i$ and $\mathbf{B}_+ = \bigcup_{i=1}^{m} \mathbf{B}_i$. Thus $|\mathbf{B}_i| = e_i$, and $|\mathbf{B}_+| = e$. For example, if $p = 3$, and $n = 16$, we have $m = 2$, the base 3 expansion of

$$n = 1 \times 3^0 + 2 \times 3^1 + 1 \times 3^2 \text{ and } e = e_1 + e_2 = 5 + 1 = 6$$

$$\begin{array}{lcccc} \mathbf{B}_0 & = & \{B_{0,1}, B_{0,2}, \ldots, B_{0,16}\} & = & \{[1], [2], \ldots, [16]\} \\ \mathbf{B}_1 & = & \{B_{1,1}, B_{1,2}, \ldots, B_{1,5}\} & = & \{[1,3], [4,6], \ldots, [13,15]\} \\ \mathbf{B}_2 & = & \{B_{2,1}\} & = & \{[1,9]\} \end{array}$$

For each element $B \in \mathbf{B}_+$ let $\sigma_B \in S_n$ be the permutation of X_n which fixes all elements of $X_n - B$ and permutes the elements of B ahead cyclically through $\frac{1}{p}$th the length of the block B. Thus, if $x \in B \in \mathbf{B}_i$ with $i \geq 1$, then $\sigma_B(x) \in B$ and $\sigma_B(x) \equiv x + p^{i-1} \pmod{p^i}$. Clearly, each σ_B has order p in S_n. Let $\Gamma_n = \{\sigma_B \mid B \in \mathbf{B}_+\}$ and let $H_n = \langle \Gamma_n \rangle$ be the subgroup of S_n generated by Γ_n. We again use the example from above:

$$\sigma_{B_{1,1}} = (1\,2\,3),\ \sigma_{B_{1,2}} = (4\,5\,6),\ \sigma_{B_{1,3}} = (7\,8\,9),$$

$$\sigma_{B_{1,4}} = (10\,11\,12),\ \sigma_{B_{1,5}} = (13\,14\,15)$$

$$\sigma_{B_{2,1}} = (1\,4\,7)(2\,5\,8)(3\,6\,9)$$

Proposition 1. *H_n is a Sylow p-subgroup of S_n.*

Proof: We will show that every element of H_n may be uniquely expressed in the form

$$\sigma = \prod_{B \in \mathbf{B}_+} \sigma_B^{a_B} \text{ with each } a_B \in \{0, \ldots, p-1\}. \tag{1}$$

For definiteness, we shall understand all such products to be expanded with the B's ordered lexicographical, so $B_{i,j}$ precedes $B_{k,h}$ (i.e. occurs to the

left of it in the product) iff either $i < k$ or else $i = k$ and $j < h$. Once this is shown we will have $|H_n| = p^e$ and the proposition will follow.

Let K denote the set of all elements $\sigma \in S_n$ which may be expressed as in (1). It is clear that $\Gamma_n \subseteq K \subseteq H_n$ and therefore, to show that $K = H_n$ (i.e. to show that every element of H_n may be expressed as in (1)), we need only show that K is closed under composition.

First note that if $B \in \mathbf{B}_+$ and $B' \in \mathbf{B}_i$, then $\sigma_B(B') \in \mathbf{B}_i$ also, and in fact $\sigma_B(B') = B'$ unless $B' \subset B$ (B' is properly contained in B). In this case, $\sigma_B(B') \neq B'$ and $\sigma_B^a(B') = B'$ iff $p|a$.

Now suppose $B \in \mathbf{B}_k$ and $B' \in \mathbf{B}_i$ with $1 \leq i \leq k$. Then either B and B' are disjoint or else $B' \subseteq B$. If B and B' are disjoint, then clearly σ_B and $\sigma_{B'}$, commute. Assume next that $B' \subseteq B$. If $B' = B$, then again σ_B and $\sigma_{B'}$ commute, being equal. So now assume that $B' \subset B$. If we were to write $\sigma_{B'}$ in disjoint cycle notation, as a product of p^{i-1} disjoint p-cycles whose constituents together comprise B', then $\sigma_B \sigma_{B'} \sigma_B^{-1}$ would equal the expression in which the constituents of these cycles have been acted upon by σ_B. The new constituents together would comprise $\sigma_B(B')$, and since σ_B maps B' onto $\sigma_B(B')$ in order preserving fashion, it is evident that $\sigma_B \sigma_{B'} \sigma_B^{-1} = \sigma_{\sigma_B(B')}$, so

$$\sigma_B \sigma_{B'} = \sigma_{\sigma_B(B')} \sigma_B \tag{2}$$

if $B' \subset B$. In fact, this holds whenever $B \in \mathbf{B}_k$ and $B \in \mathbf{B}_i$ with $1 \leq i \leq k$ since, in all other such cases, we have $\sigma_B(B') = B'$ and we already know that σ_B and $\sigma_{B'}$ commute.

Now suppose $\sigma = \prod_{B \in \mathbf{B}_+} \sigma_B^{a_B}$ and $\tau = \prod_{B \in \mathbf{B}_+} \sigma_B^{b_B}$ are given. We obtain an expression for $\sigma\tau$ by juxtaposing these, and the repeated use of (2) gives an expression $\sigma\tau = \prod_{B \in \mathbf{B}_+} \sigma_B^{c_B}$ where we may assume each $c_B \in \{0, \ldots, p-1\}$ since each σ_B has order p. Thus, $K = H_n$.

It remains only to prove uniqueness of the expressions. Let σ and τ be as in the previous paragraph and assume $\sigma = \tau$. For each $B \in \mathbf{B}_m$ we may choose some $B' \in \mathbf{B}_{m-1}$ with $B' \subset B$ and then, since σ_B is the only element of $\mathbf{B}_+$ which moves B', we have $\sigma_B^{a_B}(B') = \sigma(B') = \tau(B') = \sigma_B^{b_B}(B')$; therefore $\sigma_B^{a_B - b_B}(B') = B'$, $p \mid (a_B - b_B)$, and $a_B = b_B$. Multiplying σ and τ on the right by $\sigma_B^{-a_B}$ for all $B \in \mathbf{B}_m$ gives new expressions which we again call σ and τ and which are again equal, and so we may assume $a_B = b_B = 0$ for all $B \in \mathbf{B}_m$. This allows us (if $m \geq 2$) to argue similarly that for each $B \in \mathbf{B}_{m-1}$ we have $a_B = b_B = 0$; then (if $m \geq 3$) for $B \in \mathbf{B}_{m-2,\ldots,\text{etc.}}$, and finally for $B \in \mathbf{B}_1$. So expressions as in (1) are unique. □

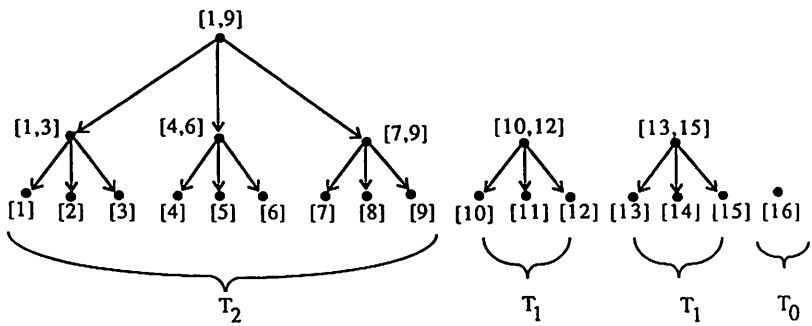
3 The number of derangements of H_n

We now look at the number of derangements in H_n. In order to aid us in doing so, we make our family $\mathbf{B}$ into a directed graph as follows. For each

block $B \in \mathbf{B}_i$ with $i \geq 1$, there are exactly p distinct elements $B' \in \mathbf{B}_{i-1}$ with $B' \subset B$, and we include an edge from B to each of these elements B'. So elements of $\mathbf{B}_+$ have out-degree p and elements of $\mathbf{B}_0$ have out-degree 0. All elements have in-degree either 0 or 1.

For our purposes, a tree will be a finite, nonempty directed graph in which a single vertex called the root has in-degree 0 and there is a unique path from the root to any other vertex. Vertices having out-degree 0 are called leaves. A forest will be a directed graph $\mathbf{F}$ which is a disjoint union of a finite number of trees, and then these trees are exactly the maximal subtrees of F. For $i \geq 0$ we let T_i denote a full p-ary tree of height i, so all paths from the root to a leaf contain i edges and every non-leaf vertex has out-degree p.

We may view X_n as a disjoint union of d_m blocks of length p^m, followed by d_{m-1} blocks of length $p^{m-1}, \ldots,$ etc., followed by d_0 blocks of length $p^0 = 1$, where d_i is the coefficient of p^i in the expansion of n. Then we see that these blocks are exactly the elements of $\mathbf{B}$ having in-degree 0, and $\mathbf{B}$ is a forest consisting of d_m trees isomorphic with T_m, d_{m-1} trees isomorphic with $T_{m-1}, \ldots,$ etc., and d_0 trees isomorphic with T_0. These are the maximal subtrees of $\mathbf{B}$ and their roots are the above blocks comprising X_n. Note also that a path in $\mathbf{B}$ is a maximal path iff it extends from the root to a leaf of one of the maximal subtrees of $\mathbf{B}$. For each element $j \in X_n$, let $\wp_j$ denote the set of blocks along the maximal path in $\mathbf{B}$ containing $[j]$, and note that we have $\wp_j = \{B \in \mathbf{B} \mid j \in B\}$. So $j \mapsto \wp_j$ gives a bijective correspondence between X_n and the set of all maximal paths in $\mathbf{B}$. Once again we use the above example to depict these concepts in what follows:



Next we consider derangements in H_n. Suppose $\sigma \in H_n$ and $j \in X_n$. If $\sigma(j) = j$, then we must have $\sigma(B) = B$ for each $B \in \wp_j$. Write $\sigma = \prod_{B \in \mathbf{B}_+} \sigma_B^{a_B}$ with each $a_B \in \{0, \ldots, p-1\}$. We claim that $\sigma(j) = j$ iff $a_B = 0$ for $B \in \wp_j^+ = \wp_j \cap \mathbf{B}_+$. Let k denote the height of the maximal subtree of $\mathbf{B}$ containing $[j]$, so $\wp_j$ consists of one element from each $\mathbf{B}_i$ with $0 \leq i \leq k$. Assume first that $\sigma(j) = j$. If $k \geq 1$ and B is the element of

$\wp_j \cap \mathbf{B}_k$, and B' the element of $\wp_j \cap \mathbf{B}_{k-1}$, then $B' = \sigma(B') = \sigma_B^{a_B}(B')$ and $a_B = 0$. Then, if $k \geq 2$, we see similarly that $a_B = 0$ if B is the element of $\wp_j \cap \mathbf{B}_{k-1}$. Similarly, $a_B = 0$ for all $B \in \wp_j^+$. Conversely, if $a_B = 0$ for all $B \in \wp_j^+$, we see that $\sigma(B) = B$ for all $B \in \wp_j$ and hence $\sigma(j) = j$, so the claim is proved.

If $\mathbf{F}$ is any forest, let $\mathbf{F}_+$ denote the set of non-leaf vertices in $\mathbf{F}$. By a coloring of $\mathbf{F}$ we shall mean an assignment of an element $a_v \in \{0, \ldots, p-1\}$ to each vertex $v \in \mathbf{F}_+$. (We note that this would usually be called a p-coloring of $\mathbf{F}_+$). By a random coloring of $\mathbf{F}$ we shall mean a coloring of $\mathbf{F}$ which is chosen probabilistically so that the a_v's are chosen independently, each according to the uniform probability distribution on $\{0, \ldots, p-1\}$. We shall call a coloring of $\mathbf{F}$ admissible iff each maximal path in $\mathbf{F}$ (from the root to a leaf of one of the maximal subtrees of $\mathbf{F}$) contains a non-leaf vertex $v \in \mathbf{F}_+$ with $a_v \neq 0$.

Proposition 2. *If h_n is the number of derangements in H_n and $f_n = \frac{h_n}{|H_n|}$, then f_n equals the probability that a random coloring of $\mathbf{B}$ is admissible, and we have $f_n = \prod_{i=0}^m q_i^{d_i}$ where $n = \sum_{i=0}^m d_i p^i$ is the base p expansion of n and $q_i = f_{p^i}$ is the probability that a random coloring of T_i is admissible.*

Proof: Writing $\sigma \in H_n$ in the form $\sigma = \prod_{B \in \mathbf{B}_+} \sigma_B^{a_B}$ with each $a_B \in \{0, \ldots, p-1\}$, we see that elements of H_n correspond to colorings of $\mathbf{B}$, and the above shows that derangements in H_n correspond to admissible colorings of $\mathbf{B}$. Choosing a coloring of $\mathbf{B}$ randomly corresponds to choosing $\sigma \in H_n$ randomly, with all $|H_n| = p^e$ elements being equally likely. Thus, f_n equals the probability that a random coloring of $\mathbf{B}$ is admissible. The maximal subtrees of $\mathbf{B}$ include d_i trees isomorphic with T_i for $0 \leq i \leq m$, and they are colored randomly and independently in a random coloring of $\mathbf{B}$. The proposition follows. □

Proposition 3. *We have $q_0 = 0$ and $q_i = (1 - \frac{1}{p}) + \frac{1}{p} \times q_{i-1}^p$ for $i \geq 1$.*

Proof: It is clear that $q_0 = 0$ since T_0 has no non-leaf vertices. Now assume $i \geq 1$. A coloring of T_i is admissible iff either the value assigned to the root is $\neq 0$ (probability $= 1 - \frac{1}{p}$) or else this value is $= 0$ (probability $= \frac{1}{p}$) but all p of the subtrees of T_i whose roots are the children of the root of T_i have admissible colorings (each probability $= q_{i-1}$). These subtrees are all isomorphic with T_{i-1} and they are colored randomly and independently in a random coloring of T_i, so the proposition follows. □

The sequence $\{f_n\}_{n=1}^\infty$

Using propositions 2 and 3, it is straightforward to compute f_n for any $n \geq 1$. To better understand how the f_n's are distributed, however, it is necessary to further investigate the q_i's.

Lemma 4. *The sequence $\{q_i\}_{i=0}^\infty$ is strictly increasing and has $\lim_{i \to \infty} q_i = 1$, and the product $\prod_{i=1}^\infty q_i = 0$ is divergent.*

Proof: We have $q_0 = 0$ and $q_i = g(q_{i-1})$ for $i \geq 1$, where $g(x) = (1 - \frac{1}{p}) + \frac{1}{p}x^p$. The properties of g we will use are that it is twice differentiable on the closed unit interval (all x such that $0 \leq x \leq 1$); $g'(x) > 0$ for $0 < x \leq 1$; there is some $c > 0$ such that $0 < g''(x) < c$ when $0 < x < 1$; and $g(1) = g'(1) = 1$. Our proof is general and would apply if g were replaced by any function having these properties.

From these properties it follows that $x < g(x) < x + \frac{c}{2}(x-1)^2$ for $0 \leq x < 1$. So $\{q_i\}_{i=1}^{\infty}$ is strictly increasing and bounded above by 1, the $\lim_{i\to\infty} q_i = L$ exists and $0 < L \leq 1$, and $g(L) = \lim_{i\to\infty} g(q_i) = \lim_{i\to\infty} q_{i+1} = L$ which forces $L = 1$.

Next let $r_i = 1 - q_i$ for $i \geq 1$. We will complete the proof by using the well known fact that $\prod_{i=1}^{\infty} q_i = 0$ iff $\sum_{i=1}^{\infty} r_i = \infty$ [2]. From the above we have $1 - g(x) > 1 - x - \frac{c}{2}(x-1)^2$ for $0 \leq x < 1$, which implies that $r_{i+1} > r_i - \frac{c}{2}r_i^2$ for $i \geq 1$. Choose any $0 < \lambda < 1$ and let $s_i = \frac{1}{i(\ln i)^{\lambda}}$, for $i \geq 2$. Applying the mean value theorem to $f(x) = \frac{1}{x(\ln x)^{\lambda}}$ on the interval $i \leq x \leq i+1$, we can see that if i is sufficiently large then $s_{i+1} < s_i - \frac{c}{2}s_i^2$. Now both $\{r_i\}$ and $\{s_i\}$ are decreasing and converge to 0. If we discard a finite number of terms from each and re-index both to begin with $i = 1$, we may assume that $s_1 < r_1$ and that for all $i \geq 1$ we have $r_i, s_i < \frac{1}{c}$ and $s_{i+1} < s_i - \frac{c}{2}s_i^2$. We may then show by induction that $s_i < r_i$ for all $i \geq 1$. For the inductive step, using the fact that $h(x) = x - \frac{c}{2}x^2$ is increasing where $0 \leq x \leq \frac{1}{c}$, we obtain $s_{i+1} < s_i - \frac{c}{2}s_i^2 < r_{i+1}$. Now $\sum s_i = \infty$ by the integral test, and hence $\sum r_i = \infty$ by the comparison test. □

Proposition 5. *The sequence $\{f_n\}_{n=1}^{\infty}$ is dense in the closed unit interval.*

Proof: Suppose $0 < \alpha < 1$. It will suffice to find an increasing sequence $\{n_k\}_{k=1}^{\infty}$ of natural numbers with $\lim_{k\to\infty} f_{n_k} = \alpha$.

Let t_1 be the smallest natural number with $q_{t_1} > \alpha$. Assuming $t_1 < \cdots < t_k$ have been chosen and $q_{t_1} \times \ldots \times q_{t_k} > \alpha$, let t_{k+1} be the smallest natural number with $t_{k+1} > t_k$ and $q_{t_1} \times \ldots \times q_{t_{k+1}} > \alpha$. This is possible because $\lim_{i\to\infty} q_i = 1$.

We claim next that $\prod_{k=1}^{\infty} q_{t_k} = \alpha$. It is clear that $\prod_{k=1}^{\infty} q_{t_k} = \beta$ exists and $\beta \geq \alpha$, since the finite subproducts are decreasing and all are $> \alpha$. So assume that $\beta > \alpha$. The $\{t_k\}_{k=1}^{\infty}$ will include every natural number t with $q_t > \frac{\alpha}{\beta}$, i.e. all sufficiently large natural numbers, and it will follow by Proposition 4 that $\prod_{k=1}^{\infty} q_{t_k} = 0$, a contradiction. So $\beta = \alpha$.

Now for each $k \geq 1$ let $n_k = \sum_{i=1}^{k} p^{t_i}$. Then $f_{n_k} = \prod_{i=1}^{k} q_{t_i}$ by Proposition 2, and hence $\lim_{k\to\infty} f_{n_k} = \alpha$. □

Finally, we show that while $\{f_n\}_{n=1}^{\infty}$ has no limit, it is Cesàro convergent to 0.

Proposition 6. *If g_N denotes the average of the first N of the f_n's, then* $\lim_{N\to\infty} g_N = 0$.

Proof: For convenience we let $f_0 = 1$, which is appropriate since the only permutation of the empty set is the function which is the empty set of ordered pairs. Since this permutation has no fixed points, we have $h_0 = |H_0| = 1$. We let $g_N = \frac{1}{N}\sum_{n=0}^{N-1} f_n$. Once we prove that $\lim_{N\to\infty} g_N = 0$, it will follow that the same would have been true had we defined g_N to be $\frac{1}{N}\sum_{n=1}^{N} f_n$.

Assume $p^m < N \leq p^{m+1}$. Then we have $\sum_{n=0}^{N-1} f_n \leq \sum_{n=0}^{p^{m+1}-1} f_n = \sum_{n=0}^{p^{m+1}-1} q_0^{d_0}\times\ldots\times q_m^{d_m}$ (where the d_i's depend on n, each $d_i \in \{0,\ldots,p-1\}$, and $n = \sum_{i=0}^{m} d_i p^i$) $= \sum_{d_0,\ldots,d_m=0}^{p-1} q_0^{d_0} \times \ldots \times q_m^{d_m} = \prod_{i=0}^{m}(\sum_{d_i=0}^{p-1} q_i^{d_i}) = \prod_{i=0}^{m} \frac{1-q_i^p}{1-q_i}$. The recurrence in Proposition 3 implies that $1-q_i^p = p(1-q_{i+1})$, so $\sum_{n=0}^{N-1} f_n \leq \prod_{i=0}^{m} \frac{p(1-q_{i+1})}{1-q_i} = p^{m+1}(1-q_{m+1})$. We also have $\frac{1}{N} < \frac{1}{p^m}$, so that $g_N = \frac{1}{N}\sum_{n=0}^{N-1} f_n < p(1-q_{m+1})$. Since $\lim_{i\to\infty} q_i = 1$ by Lemma 4, it follows that $\lim_{N\to\infty} g_N = 0$. □

References

[1] R. Alperin, Derangements, preprint, 1991.

[2] E.C. Titchmarsh, *The Theory of Functions*, Oxford University Press, London, 1939, (15).