

ON ALMOST-COMMUTING PERMUTATIONS

DANIEL GORENSTEIN, REUBEN SANDLER AND W. H. MILLS

Suppose A and B are two permutations on a finite set X which commute on almost all of the points of X . Under what circumstances can we conclude that B is approximately equal to a permutation which actually commutes with A ? The answer to this question depends strongly upon the order of the centralizer, $C(A)$, of A in the symmetric group on X ; and this varies greatly according to the cycle structure of A , being comparatively small when A is either a product of few disjoint cycles or a product of a large number of disjoint cycles of different lengths and being comparatively large when A is a product of many disjoint cycles, all of the same length. We shall show by example that when the order of $C(A)$ is small there may exist a permutation B which commutes with A "almost everywhere" yet is not approximated by any element of $C(A)$. On the other hand, when A is a product of many disjoint cycles of the same length, we shall see that for any such permutation B , there must exist a permutation in $C(A)$ which agrees closely with B .

It is clear that if B is a permutation leaving fixed almost all points of X , then no matter what permutation A is given, B will commute with A on almost all points of X , and at the same time B can be closely approximated by an element of $C(A)$ —namely, the identity. However, the examples we shall give will show that only when all (or nearly all) of the cycles of A are of the same length can we hope to approximate *every* B which nearly commutes with A by an element in $C(A)$. Accordingly, the bulk of this paper will be taken up with the study of the case in which A is a product of many disjoint cycles, all of the same length.

1. In order to get a satisfactory notation and a more compact way of discussing the problem, we begin by making the symmetric group $S_N(X)$ on the space X into a metric space. Here N denotes the cardinality of X , and it is to be understood that N is finite. Define, for any A in $S_N(X)$,

$$(1) \quad \|A\| = \frac{N - f_A}{N}$$

where f_A is the number of fixed points of A on X . Now define the distance $d(A, B)$ between two elements A and B of $S_N(X)$ to be

$$(2) \quad d(A, B) = \|AB^{-1}\|.$$

Received December 2, 1961,

Under these definitions, the identity is the only permutation of norm 0, every permutation has norm ≤ 1 , and a permutation has norm p if and only if it moves pN points of X . In particular, the permutations A and B commute if and only if $\|[A, B]\| = 0$, or equivalently, if and only if $d(AB, BA) = 0$.

In order to see that these definitions make $S_N(X)$ into a metric space, we need only verify the triangle inequality, since the other properties are trivial. But the points of X displaced by AB are clearly among those which are displaced by either A or B . Hence $N - f_{AB} \leq (N - f_A) + (N - f_B)$ and consequently $\|AB\| \leq \|A\| + \|B\|$. We thus have the following lemma.

LEMMA 1. *With the norm defined above, $S_N(X)$ forms a metric space.*

When no restriction is placed upon the cycle structure of A , we have the following result:

PROPOSITION 1. For any $\varepsilon > 0$, there exists an integer N and permutations A and B in $S_N(X)$ such that $\|[A, B]\| < \varepsilon$ and such that $d(B, D) = 1$ for every D in $C(A)$.

Proof. We shall give two examples of permutations A and B which satisfy the conditions of the proposition; in the first, A will be a product of cycles of relatively prime lengths, and in the second, a product of cycles of lengths n and $2n$.

EXAMPLE 1. Let $X = \{1, 2, \dots, N\}$, where $N = 2n > 4/\varepsilon$. Let A be the permutation

$$(1\ 2 \cdots n - 1)(n)(n + 1\ n + 2 \cdots 2n)$$

and B the permutation $xB = x + n$ if $x \leq n$, and $xB = x - n$ if $x > n$. By direct verification, we find that A and B commute except on the points $n - 1, n, 2n - 1, 2n$. Thus $f_{[A, B]} = N - 4$ and hence $\|[A, B]\| = 4/N < \varepsilon$.

On the other hand, any element D of $C(A)$ must map each cycle of A into itself, since these cycles are of different lengths. But, for any x in X , xB and x lie in distinct cycles of A . It follows that for any D in $C(A)$, BD^{-1} displaces every point of X and hence that $d(B, D) = 1$.

EXAMPLE 2. Let $X = \{1, 2, \dots, N\}$, where $N = 4nm$ and $n > 1/\varepsilon$. Let A be the permutation with m cycles of length $2n$ and $2m$ cycles

of length n , defined as follows:

$$(1\ 2 \cdots 2n)(2n + 1 \cdots 4n) \cdots (2n(m - 1) + 1 \cdots 2nm) \\ (2nm + 1 \cdots 2nm + n)(2nm + n + 1 \cdots 2nm + 2n) \cdots \\ (4nm - n + 1 \cdots 4nm).$$

Let B be the permutation $xB = x + 2nm$ if $x \leq 2nm$, and $xB = x - 2nm$ if $x > 2nm$.

Again, by direct computation, we find that A and B commute on all points x of X except when $x \equiv 0 \pmod{n}$. Thus $f_{[A,B]} = 4nm - 4m$ and hence $\|[A, B]\| = 1/n < \epsilon$. On the other hand, if $D \in C(A)$, D must permute the cycles of A of length n among themselves and must permute the cycles of A of length $2n$ among themselves. But if x is in a cycle of length n , then xB is in a cycle of length $2n$, and vice versa. It follows that BD^{-1} displaces every point of X and hence that $d(B, D) = 1$, for any D in $C(A)$.

2. The two examples given in Proposition 1 indicate that unless severe restrictions are placed on the cycle structure of A , the fact that B comes very close to commuting with A does not necessarily imply that B can be approximated by an element in $C(A)$. In fact, it seems that unless A consists almost entirely of cycles of the same length, little can be said in general of the relation between $\|[A, B]\|$ and the distance from B to $C(A)$.

In order to be able to make as exact statements as possible, we shall assume in the balance of the paper that A is the product of m disjoint cycles, each of length n . In this case our statements about the distance from B to $C(A)$ will depend only upon $\|[A, B]\|$ and n .

We may take $X = \{1, 2, \dots, N\}$, where now $N = nm$. Let x, k be integers such that $1 \leq x \leq N$, $0 \leq k \leq n$, and write $x = in + r$, where $1 \leq r \leq n$. We shall adopt the following notation:

$$(3) \quad \overline{x+k} = in + s, \text{ where } 1 \leq s \leq n \text{ and } s \equiv r + k \pmod{n}.$$

Without loss of generality we may assume that A is the mapping

$$(4) \quad xA = \overline{x+1}, \quad x \in X.$$

We shall say that B in $S_N(X)$ transforms the cycle a of A into the cycle a' if, for some x in a , xB is in a' and

$$(5) \quad \overline{(x+k)B} = \overline{xB+k}, \quad k = 0, 1, \dots, n-1.$$

We shall write $(a)B = a'$ if B transforms a into a' . We shall also say that B commutes with A on a cycle a if it commutes with A on each point of a .

LEMMA 2. (a) *A permutation B commutes with A on a cycle a if and only if B transforms a into a cycle a'.*

(b) *if B commutes with A on n - 1 points of a cycle a, then B commutes with A on a.*

(c) *If B transforms r cycles of A into cycles of A, there exists an element D in C(A) which agrees with B on these r cycles.*

Proof. For A and B to commute on a point x of X we must have $xBA = xAB$, and hence

$$(6) \quad \overline{xB + 1} = \overline{(x + 1)B}.$$

Suppose $(a)B = a'$; then (6) follows at once from (5) for any x in a. Conversely if (6) holds for all x in a, (5) follows at once by induction on k.

To prove (b), suppose B and A commute on $x, \overline{x + 1}, \dots, \overline{x + n - 2}$. Again by induction on k, (5) holds for $k = 0, 1, \dots, n - 2$. In particular, $\overline{(x + n - 2)B} = \overline{xB + n - 2}$. Now using (6) with x replaced by $\overline{x + n - 2}$, we obtain

$$\begin{aligned} \overline{(x + n - 1)B} &= \overline{\overline{(x + n - 2)B} + 1} \\ &= \overline{\overline{xB + n - 2} + 1} = \overline{xB + n - 1}. \end{aligned}$$

Thus (5) holds for all k, and hence A and B commute on a by part (a).

Finally suppose B transforms the cycles a_1, \dots, a_r into the cycles a'_1, \dots, a'_r . Denote by a'_{r+1}, \dots, a'_m the remaining cycles of A. Let D be a permutation which agrees with B on a_1, \dots, a_r and transforms a_i into $a'_i, i = r + 1, \dots, m$. By (a) D is in C(A).

3. We shall now begin the analysis of the relationship between $\| [A, B] \|$ and the minimum distance from B to C(A), under the assumption that A is the product of n-cycles. We shall denote this minimum distance by $d_A(B)$. Thus

$$(7) \quad d_A(B) = \min_{D \in C(A)} d(B, D).$$

Then following estimate for $d_A(B)$ is easily obtained.

PROPOSITION 2. For any B in $S_N(X)$,

$$d_A(B) \leq \frac{n \| [A, B] \|}{2}.$$

Proof. If $\| [A, B] \| \geq 2/n$, the proposition is vacuously true since $d_A(B) \leq 1$. Hence we may assume that $\| [A, B] \| < 2/n$.

Now $N = nm$, where m is the number of cycles in A . It suffices to show that B transforms at least

$$m - \frac{N \|[A, B]\|}{2}$$

cycles of A into cycles of A . For then by Lemma 2(c) we can find an element D in $C(A)$ which agrees with B on these cycles and hence on at least

$$N - \frac{nN}{2} \cdot \|[A, B]\|$$

points of X . It follows that

$$d(B, D) \leq \frac{n \|[A, B]\|}{2}.$$

By the definition of $\|[A, B]\|$, $N \cdot \|[A, B]\|$ is the number of points displaced by $[A, B]$ and hence on which A and B do not commute. But by Lemma 2(b) any cycle of A which is not transformed by B into a cycle of A contains at least 2 points on which A and B do not commute. Thus there are at most

$$\frac{N \|[A, B]\|}{2}$$

cycles of A which are not transformed by B into cycles of A , and hence B transforms at least

$$m - \frac{N \|[A, B]\|}{2}$$

cycles of A into cycles of A .

Proposition 2 gives an upper bound for $d_A(B)$, which depends only upon $\|[A, B]\|$ (and n), but not upon the particular structure of B . Our main concern in the paper will be in improving this upper bound. The next proposition shows the limit to which this estimate can be improved.

PROPOSITION 3. If A contains at least two distinct cycles, then there exists a permutation B in $S_N(X)$ such that

$$d_A(B) = \frac{n \|[A, B]\|}{4}$$

when n is even, and such that

$$d_A(B) = \frac{n-1}{4} \|[A, B]\|$$

when n is odd. Furthermore for any $\varepsilon > 0$, N and B can be chosen so that $\| [A, B] \| < \varepsilon$.

Proof. Assume first that n is even. Set $m = m_1 + m_2$, where $m_1 \geq 0$ and $m_2 \geq 2$. Define the permutation B as follows: $xB = x$ if $1 \leq x \leq nm_1$; if $x > nm_1$, write $x = in + k$ where $1 \leq k \leq n$, and define $xB = x$ if $k \leq n/2$, $xB = x + n$ if $i \neq m - 1$ and $k > n/2$, and $xB = nm_1 + k$ if $i = m - 1$ and $k > n/2$.

Thus B leaves the first m_1 cycles of A pointwise fixed, one half of each of the remaining m_2 cycles pointwise fixed, and permutes the other halves of these m_2 cycles cyclically. From its definition, we see that B commutes with A except on the points $x > nm_1$ for which $x \equiv 0 \pmod{n/2}$. Thus

$$(8) \quad \| [A, B] \| = \frac{2m_2}{N} .$$

Since $N = n(m_1 + m_2)$, $2m_2/N$ can be made arbitrarily small by making m_1 sufficiently large. Thus, to prove the proposition, we have only to show that

$$d_A(B) = \frac{n \| [A, B] \|}{4} .$$

Observe, first of all, that the identity, I , is in $C(A)$ and agrees with B on

$$nm_1 + \frac{nm_2}{2}$$

points of X , whence

$$(9) \quad d(I, B) = \frac{N - nm_1 - \frac{nm_2}{2}}{N} = \frac{nm_2}{2N} = \frac{n}{4} \| [A, B] \| .$$

On the other hand, by Lemma 2, any D in $C(A)$ must transform each cycle a_i of A into some other cycle a_j . Since B transforms the two halves of the cycles a_i into distinct cycles of A , $m_1 \leq i \leq -1$, D and B can agree on at most half of the nm_2 points in these cycles. Hence DB^{-1} displaces at least $nm_2/2$ points of X , which implies that

$$d(D, B) \geq \frac{nm_2}{2N} = \frac{n}{4} \| [A, B] \|$$

for any D in $C(A)$.

When n is odd, the construction of B is entirely analogous.

4: If we set

$$d_A = \max_{\substack{B \in S_N(X) \\ B \in C(A)}} \frac{d_A(B)}{\|[A, B]\| n} ,$$

then d_A is a measure of the extent to which every permutation in $S_N(X)$ can be approximated by elements in $C(A)$. Propositions 2 and 3 show that

$$(10) \quad \frac{1}{4} \leq d_A \leq \frac{1}{2} \text{ or } \frac{n-1}{4n} \leq d_A \leq \frac{1}{2}$$

according as n is even or odd.

In the balance of the paper we shall sharpen these inequalities by lowering the upper bound for d_A . Our next result will show that in considering this problem, we may restrict our attention to those cycles of A on which B commutes with A on exactly $n, n - 2$, or $n - 3$ points. Let U_B, V_B, W_B be the set of points in those cycles of A on which B commutes with A on $n, n - 2$, and $n - 3$ points respectively; and let $u_B = |U_B|, v_B = |V_B|, w_B = |W_B|$.

THEOREM 1. *Suppose there exists an element D in $C(A)$ which agrees with B on at least $u_B + (1/2)v_B + (1/3)w_B$ points of X . Then*

$$d_A(B) \leq \|[A, B]\| \frac{n}{4} .$$

Proof. For simplicity of notation, we drop the subscript B , and define

$$(11) \quad t = N - u - v - w .$$

Thus t is the number of points in those cycles of A on which A and B commute on no more than $n - 4$ points. Then by definition of u, v, w, t , we have

$$(12) \quad u + \frac{n-2}{n}v + \frac{n-3}{n}w + \frac{n-4}{n}t \geq f_{[A, B]} .$$

Now, by hypothesis,

$$(13) \quad d(B, D) \leq \frac{N - \left(u + \frac{1}{2}v + \frac{1}{3}w\right)}{N} = \frac{\frac{1}{2}v + \frac{2}{3}w + t}{N} .$$

We must show that

$$(14) \quad \frac{\frac{1}{2}v + \frac{2}{3}w + t}{N} \leq \frac{n}{4} \|[A, B]\| .$$

But using (1), we can rewrite (14) as:

$$(15) \quad f_{[A,B]} \leq u + \frac{n-2}{n}v + \left(1 - \frac{8}{3n}\right)w + \left(\frac{n-4}{n}\right)t.$$

Since (15) is an immediate consequence of (12), the theorem follows.

5. In this section, we prove that $d_A \leq 1/4$, by proving that for any B in $S_N(X)$, there exists a permutation D in $C(A)$ which satisfies the conditions of Theorem 1.

To treat our problem, we need an additional concept: By a *block* of a cycle a of A , we shall mean a maximal sequence $\overline{x, x+1, \dots, x+r-1}$ of points of a such that A and B commute on every point of the sequence except $\overline{x+r-1}$. The integer r will denote the *length* of the block. According to the definition, if A and B commute on every point of a then a contains no blocks. When B and A do not commute on every point of a , we have the following obvious lemma:

LEMMA 3. *If A and B commute on exactly $n-k$ points of a cycle a of A , $k > 0$, then A contains exactly k blocks, the sum of whose lengths is n .*

Thus when a cycle a of A lies in V_B , a consists of 2 blocks which we denote by p_1, p_2 ; and when a lies in W_B , a consists of 3 blocks which we denote by q_1, q_2, q_3 . We define $|p_j|, |q_j|$ to be the lengths of p_j, q_j , respectively. Furthermore we order the blocks so that $|p_1| \geq |p_2|$ and $|q_1| \geq |q_2| \geq |q_3|$. Since $|p_1| + |p_2| = n$,

$$(16) \quad |p_1| \geq \frac{n}{2}$$

and likewise

$$(17) \quad |q_1| \geq \frac{n}{3}.$$

Let $\overline{x, x+1, \dots, x+r-1}$ be a block contained in a cycle a . If $xB = y$, then, it follows from (6) as in the proof of Lemma 2, that

$$(18) \quad (\overline{x+k})B = \overline{y+k}, \quad 0 \leq k \leq r-1;$$

and

$$(19) \quad (\overline{x+r})B \neq \overline{y+r}.$$

Thus the image of the block is a consecutive sequence of points in a cycle a' . It follows that there exist permutations which transform a

into a' and agree with B on the block $b = \{x, \overline{x+1}, \dots, \overline{x+r-1}\}$. In fact, any D in $C(A)$ for which $xD = y$ has this property. If D is such a permutation, we shall write simply $(a)D = a'$; $(b)D = (b)B$.

From this fact, we easily derive the following lemma:

LEMMA 4. *Let a_1, \dots, a_k be distinct cycles of A containing the blocks b_1, \dots, b_k respectively. If the images of b_i under B lie in distinct cycles a'_i of A , $i = 1, 2, \dots, k$, then there exist permutations D in $C(A)$ such that $(a_i)D = a'_i$; $(b_i)D = (b_i)B$, $i = 1, 2, \dots, k$.*

We are now in a position to prove the following result:

THEOREM 2. *Given any B in $S_N(X)$, there exists an element D in $C(A)$ which agrees with B on at least*

$$u_B + \frac{1}{2}v_B + \frac{1}{3}w_B$$

points of X .

Proof. Let a_1, a_2, \dots, a_m be the cycles of A . For any $i, j, 1 \leq i, j \leq m$, let b_{ij} be the maximal number of elements of a_i on which a permutation D in $C(A)$ mapping a_i into a_j can agree with B . Thus if B transforms a_i into a_j , $b_{ij} = n$. If $(a_i)B \cap a_j = \phi$, then $b_{ij} = 0$. Now, to any $m \times m$ permutation matrix (e_{ij}) there corresponds a permutation D in $C(A)$ which agrees with B on

$$(20) \quad \sum_{i,j} e_{ij}b_{ij}$$

points, where D is defined to transform a_i into a_j if $e_{ij} = 1$, and to map a_i so as to agree with B on b_{ij} points.

We wish to show

$$(21) \quad \max_{(e_{ij})} \sum e_{ij}b_{ij} \geq u + \frac{1}{2}v + \frac{1}{3}w,$$

where (e_{ij}) ranges over all permutation matrices. To do this, consider the set of all real $m \times m$ matrices (x_{ij}) such that

$$(22) \quad x_{ij} \geq 0; \quad 1 \leq i, j \leq m$$

$$(23) \quad \sum_i x_{ij} = 1; \quad 1 \leq j \leq m$$

$$(24) \quad \sum_j x_{ij} = 1; \quad 1 \leq i \leq m.$$

This is the set of doubly stochastic matrices and is a convex, bounded set whose vertices consist of exactly the permutation matrices (see [1],

pp. 132-3).

The following lemma will be useful in proving the theorem.

LEMMA 5. *If (x_{ij}) is any doubly stochastic matrix, then there exists a permutation matrix (e_{ij}) such that*

$$(25) \quad \sum_{i,j} e_{ij} b_{ij} \geq \sum_{i,j} x_{ij} b_{ij} .$$

Proof. See [1], p. 134.

If we can now demonstrate a doubly stochastic matrix such that

$$(26) \quad \sum_{i,j} x_{ij} b_{ij} \geq u + \frac{1}{2}v + \frac{1}{3}w ,$$

we will clearly be finished since, by Lemma 5, there must then be some permutation matrix (e_{ij}) such that

$$\sum_{i,j} e_{ij} b_{ij} \geq u + \frac{1}{2}v + \frac{1}{3}w ,$$

and this permutation matrix will yield the desired mapping D .

To find a matrix satisfying (26), define

$$(27) \quad x_{ij} = \frac{n_{ij}}{n} ,$$

where n_{ij} is the number of points of a_i which B maps into a_j . The matrix (x_{ij}) is clearly doubly stochastic, so we must show that (26) holds. But if $a_i \subseteq U_B$, then

$$\sum_j x_{ij} b_{ij} = n ,$$

since $(a_i)B = a_{j_1}$ for some j_1 . If $a_i \subseteq V_B$, there exist indices j_1 and j_2 such that $(p_1)B \subset a_{j_1}$ and $(p_2)B \subset a_{j_2}$. Note that $j_1 \neq j_2$, or else a_i would be transformed by B into a_{j_1} . In this case, then,

$$\sum_j x_{ij} b_{ij} = \frac{|p_1|^2}{n} + \frac{|p_2|^2}{n} \geq \frac{n}{2}$$

(remember $|p_1| + |p_2| = n$).

Finally, when $a_i \subseteq W_B$, one of three things can happen:

- (a) q_1, q_2, q_3 can be mapped by B into three distinct cycles of A .
- (b) q_1, q_2, q_3 can be mapped by B into only two cycles of A ,
- (c) q_1, q_2, q_3 can be mapped into one cycle of A .

In the first case,

$$\sum_j x_{ij} b_{ij} = \sum_{k=1}^3 \frac{|q_k|^2}{n} .$$

In the second case,

$$\sum_j x_{ij} b_{ij} = \frac{|q_{k_1}|^2}{n} + \frac{(|q_{k_2}| + |q_{k_3}|)}{n} |q_{k_2}|$$

where $|q_{k_2}| \geq |q_{k_3}|$.

Finally, in case *c*,

$$\sum_j x_{ij} b_{ij} = |q_1| \frac{(|q_1| + |q_2| + |q_3|)}{n},$$

where $|q_1| \geq |q_2|, |q_3|$.

Since $|q_1| + |q_2| + |q_3| = n$, it follows at once in all three cases that

$$\sum_j x_{ij} b_{ij} \geq \frac{n}{3}.$$

We have thus demonstrated the existence of a doubly stochastic matrix (x_{ij}) with the property

$$\sum_{i,j} x_{ij} b_{ij} \geq u + \frac{1}{2}v + \frac{1}{3}w.$$

Together with Lemma 5, this completes the proof of the theorem.

As an immediate corollary of Theorems 1 and 2, we obtain our main result:

THEOREM 3. *Let A contain at least two distinct cycles. If n is even, $d_A = 1/4$. If n is odd,*

$$\frac{n-1}{4n} \leq d_A \leq \frac{1}{4}.$$

REFERENCE

1. S. Karlin, *Mathematical Methods and Theory in Games, Programming and Economics*, vol. I (1959).

CLARK UNIVERSITY
INSTITUTE FOR DEFENSE ANALYSES
YALE UNIVERSITY

