This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2016.2620160, IEEE Transactions on Information Theory

Perfect Snake-in-the-Box Codes for Rank Modulation

Alexander E. Holroyd

***Abstract*—For odd n, the alternating group on n elements is generated by the permutations that jump an element from an odd position to position 1. We prove Hamiltonicity of the associated directed Cayley graph for all odd $n \neq 5$. (A result of Rankin implies that the graph is not Hamiltonian for $n = 5$.) This solves a problem arising in rank modulation schemes for flash memory. Our result disproves a conjecture of Horovitz and Etzion, and proves another conjecture of Yehezkeally and Schwartz.**

***Index Terms*—Hamiltonian cycle, Cayley graph, snake-in-the-box, Gray code, rank modulation.**

I. Introduction

The following questions are motivated by applications involving flash memory. Let S_n be the **symmetric group** of permutations $\pi = [\pi(1), \ldots, \pi(n)]$ of $[n] := \{1, \ldots, n\}$, with composition defined by $(\pi\rho)(i) = \pi(\rho(i))$. For $2 \leq k \leq n$ let

$$\tau_k := \big[k, \;\; 1, 2, \ldots, k-1, \;\; k+1, \ldots, n\big] \in S_n$$

be the permutation that jumps element k to position 1 while shifting elements $1, 2, \ldots, k-1$ right by one place. Let $\mathcal{S}_n$ be the **directed Cayley graph** of S_n with generators $\tau_2, \ldots, \tau_n$, i.e. the directed graph with vertex set S_n and a directed edge, labelled τ_i, from π to $\pi\tau_i$ for each $\pi \in S_n$ and each $i = 2, \ldots, n$.

We are concerned with self-avoiding directed cycles (henceforth referred to simply as **cycles** except where explicitly stated otherwise) in $\mathcal{S}_n$. (A cycle is self-avoiding if it visits each vertex at most once). In applications to flash memory, a permutation represents the relative ranking of charges stored in n cells. Applying τ_i corresponds to the operation of increasing the ith charge to make it the largest, and a cycle is a schedule for visiting a set of distinct charge rankings via such operations. Schemes of this kind were originally proposed in [1].

One is interested in maximizing the length of such a cycle, since this maximizes the information that can be stored. It is known that $\mathcal{S}_n$ has a directed **Hamiltonian** cycle, i.e. one that includes *every* permutation exactly once; see e.g. [1]–[3]. However, for the application it is desirable that the cycle should not contain any two permutations that are within a certain fixed distance r of each other, with respect to some metric d on S_n. The motivation is to avoid errors arising from one permutation being mistaken for another [1], [4]. The problem of maximizing cycle length for given r, d combines notions of Gray codes [5] and error-detecting/correcting codes [6], and is sometimes known as a snake-in-the-box problem. (This term has its origins in the study of analogous questions involving binary strings as opposed to permutations; see e.g. [7]).

The main result of this article is that, in the case that has received most attention (described immediately below) there is a cycle that is **perfect**, i.e. that has the maximum size even among arbitrary sets of permutations satisfying the distance constraint.

More precisely, our focus is following case considered in [8]–[10]. Let $r = 1$ and let d be the **Kendall tau** metric [11], which is defined by setting $d(\pi, \sigma)$ to be the inversion number of $\pi^{-1}\sigma$, i.e. the minimum number of elementary transpositions needed to get from π to σ. (The ith elementary transposition swaps the permutation elements in positions i and $i+1$, where $1 \leq i \leq n-1$). Thus, the cycle is not allowed to contain any two permutations that are related by a single elementary transposition. The primary object of interest is the maximum possible length M_n of such a directed cycle in $\mathcal{S}_n$.

It is easy to see that $M_n \leq n!/2$. Indeed, any set of permutations satisfying the above distance constraint includes at most one from the pair $\{\pi, \pi\tau_2\}$ for every π, but these pairs partition $\mathcal{S}_n$. To get a long cycle, an obvious approach is to restrict to the **alternating group** A_n of all even permutations. Since an elementary transposition changes the parity of a permutation, this guarantees that the distance condition is satisfied. The generator τ_k lies in A_n if and only if k is odd. Therefore, if n is odd, this approach reduces to the problem of finding a maximum directed cycle in the directed Cayley graph $\mathcal{A}_n$ of A_n with generators $\tau_3, \tau_5, \ldots, \tau_n$. Yehezkeally and Schwartz [8] conjectured that for odd n the maximum cycle length M_n is attained by a cycle of this type; our result will imply this. (For even n this approach is less useful, since without using τ_n we can access only permutations that fix n.) As in [8]–[10], we focus mainly on odd n.

For small odd n, it is not too difficult to find cycles in $\mathcal{A}_n$ with length reasonably close to the upper bound $n!/2$, by ad-hoc methods. Finding systematic approaches that work for all n is more challenging. Moreover, getting all the way to $n!/2$ apparently involves a fundamental obstacle, but we will show how it can be overcome.

Specifically, it is obvious that $M_3 = 3!/2 = 3$. For general odd $n \geq 5$, Yehezkeally and Schwartz [8] proved the inductive bound $M_n \geq n(n-2)M_{n-2}$, leading to $M_n \geq \Omega(n!/\sqrt{n})$ asymptotically. They also showed by computer search that $M_5 = 5!/2 - 3 = 57$. Horowitz and Etzion [9] improved the inductive bound to $M_n \geq (n^2 - n - 1)M_{n-2}$, giving

Microsoft Research, Redmond, USA. holroyd@microsoft.com.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2016.2620160, IEEE Transactions on Information Theory

$M_n = \Omega(n!)$. They also proposed an approach for constructing a longer cycle of length $n!/2 - n + 2 (= (1 - o(1))n!/2)$, and showed by computer search that it works for $n = 7$ and $n = 9$. They conjectured that this bound is optimal for all odd n. Zhang and Ge [10] proved that the scheme of [9] works for all odd n, establishing $M_n \geq n!/2 - n + 2$, and proposed another scheme aimed at improving the bound by 2 to $n!/2-n+4$. Zhang and Ge proved that their scheme works for $n = 7$, disproving the conjecture of [9] in this case, but were unable to prove it for general odd n.

The obvious central question here is whether there exists a perfect cycle, i.e. one of length $n!/2$, for any odd $n > 3$. As mentioned above, Horovitz and Etzion [9] conjectured a negative answer for all such n, while the authors of [8], [10] also speculate that the answer is negative. We prove a *positive* answer for $n \neq 5$.

Theorem 1. *For all odd $n \geq 7$, there exists a directed Hamiltonian cycle of the directed Cayley graph $\mathcal{A}_n$ of the alternating group A_n with generators $\tau_3, \tau_5, \ldots, \tau_n$. Thus, $M_n = n!/2$.*

Besides being the first of optimal length, our cycle has a somewhat simpler structure than those in [9], [10]. It may in principle be described via an explicit rule that specifies which generator should immediately follow each permutation π, as a function of π. (See [2], [12], [13] for other cycles of that can be described in this way). While the improvement from $n!/2 - n + 2$ to $n!/2$ is in itself unlikely to be important for applications, our methods are quite general, and it is hoped that they will prove useful for related problems.

We briefly discuss even n. Clearly, one approach is to simply leave the last element of the permutation fixed, and use a cycle in $\mathcal{A}_{n-1}$, which gives $M_n \geq M_{n-1}$ for even n. Horovitz and Etzion [9] asked for a proof or disproof that this is optimal. We expect that one can do much better. We believe that $M_n \geq (1 - o(1))n!/2$ asymptotically as $n \to \infty$ (an n-fold improvement over $(n-1)!/2$), and perhaps even $M_n \geq n!/2 - O(n^2)$. We outline a possible approach to showing bounds of this sort, although it appears that a full proof for general even n would be rather messy. When $n = 6$ we use this approach to show $M_6 \geq 315$, improving the bound $M_6 \geq 57$ of [9] by more than a factor of 5.

Hamiltonian cycles of Cayley graphs have been extensively studied, although general results are relatively few. See e.g. [3], [14]–[16] for surveys. In particular, it is unknown whether every *undirected* Cayley graph is Hamiltonian. Our key construction (described in the next section) appears to be novel in the context of this literature also.

Central to our proof are techniques having their origins in change ringing (English-style church bell ringing). Change ringing is also concerned with self-avoiding cycles in Cayley graphs of permutations groups (with a permutation representing an order in which bells are rung), and change ringers discovered key aspects of group theory considerably before mathematicians did – see e.g. [17]–[20]. As we shall see, the fact that $\mathcal{A}_5$ has no Hamiltonian cycle (so that we have the strict inequality $M_5 < 5!/2$) follows from a theorem of Rankin [21], [22] that was originally motivated by change ringing.

II. Breaking the parity barrier

In this section we explain the key obstruction that frustrated the previous attempts at a Hamiltonian cycle of $\mathcal{A}_n$ in [8]–[10]. We then explain how it can be overcome. We will then use these ideas to prove Theorem 1 in Sections III and IV.

By a **cycle cover** of a directed Cayley graph we mean a set of self-avoiding directed cycles whose vertex sets partition the vertex set of the graph. A cycle or a cycle cover can be specified in several equivalent ways: we can list the vertices or edges encountered by a cycle in order, or we can specify a starting vertex of a cycle and list the generators it uses in order, or we can specify which generator immediately follows each vertex – i.e. the label of the unique outgoing edge that belongs to the cycle or cycle cover. It will be useful to switch between these alternative viewpoints.

A standard approach to constructing a Hamiltonian cycle is to start with a cycle cover, and then successively make local modifications that unite several cycles into one, until we have a single cycle. (See [2], [8]–[10], [12], [14]–[16], [19], [23]–[25] for examples.) However, in $\mathcal{A}_n$ and many other natural cases, there is a serious obstacle involving parity, as we explain next.

The **order** $\mathrm{order}(g)$ of a group element g is the smallest $t \geq 1$ such that $g^t = \mathrm{id}$, where id is the identity. In our case, let τ_k, τ_ℓ be two distinct generators of $\mathcal{A}_n$, and observe that their ratio $\rho := \tau_\ell \tau_k^{-1}$ is simply the permutation that jumps element ℓ to position k while shifting the intervening elements by 1. For example, when $n = 9$ we have $\tau_9 = [912345678]$ and $\tau_7^{-1} = [234567189]$, so $\tau_9\tau_7^{-1} = [123456978]$ (element 9 jumps first to position 1 and then back to position 7). In general, the ratio ρ has order $q := |k - \ell| + 1$, which is odd. In the example, $q = 3$.

The fact that $\mathrm{order}(\rho) = q$ corresponds to the fact that in the Cayley graph $\mathcal{A}_n$, starting from any vertex, there is a cycle of length $2q$ consisting of directed edges oriented in *alternating* directions and with alternating labels τ_ℓ and τ_k. Consider one such alternating cycle Q, and suppose that we have a cycle cover that includes all q of the τ_k-edges of Q. Consequently, it includes none of the τ_ℓ-edges of Q (since it must include only one outgoing edge from each vertex). An example is the cycle cover that uses the outgoing τ_k-edge from every vertex of $\mathcal{A}_n$. Then we may modify the cycle cover as follows: delete all the τ_k-edges of Q, and add all the τ_ℓ-edges of Q. This results in a new cycle cover, because each vertex of the graph still has exactly one incoming edge and one outgoing edge present.

Suppose moreover that all the τ_k-edges of Q lay in distinct cycles in the original cycle cover. Then the effect of the modification is precisely to unite these q cycles into one new cycle (having the same vertices). The new cycle alternately traverses the new τ_ℓ-edges and the remaining parts of the q original cycles. All other cycles of the cycle cover are unaffected. See Fig. 1 (left) for the case $(k, \ell) = (n-2, n)$ (with $q = 3$), and Fig. 1 (right) for the permutations at the vertices of the alternating cycle Q.

A modification of the above type reduces the total number of cycles in the cycle cover by $q - 1$, and therefore, since q is odd, it does not change the *parity of the total number of*

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2016.2620160, IEEE Transactions on Information Theory

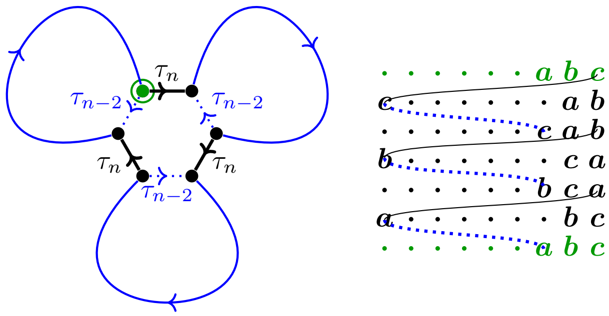

Fig. 1. *Left:* linking 3 cycles by replacing generator τ_{n-2} with generator τ_n in 3 places. We start with the 3 thin blue cycles, each of which comprises a dotted edge labeled with generator τ_{n-2}, and a curved arc that represents the remaining part of the cycle. We delete the dotted edges and replace them with the thick solid black edges (labelled τ_n), to obtain one (solid) cycle, containing the same vertices as the original 3 cycles. *Right:* the permutations at the six vertices that are marked with solid discs in the left picture. The permutation at the (green) circled vertex is $[\ldots\ldots, a, b, c]$, where $a, b, c \in [n]$, and the permutations are listed in clockwise order around the inner hexagon starting and finishing there. The ellipsis $\cdots\cdots$ represents a sequence of $n-3$ distinct elements of $[n]$, the same sequence everywhere it occurs. A solid black curve indicates that the ratio between the two successive permutations is τ_n (so that an element jumps from position n to 1), while a dotted blue curve indicates τ_{n-2}^{-1} (with a jump from 1 to $n-2$).

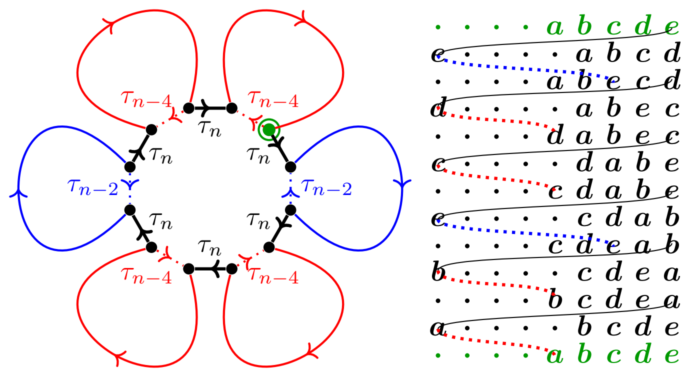

Fig. 2. The key construction. *Left:* replacing a suitable combination of generators τ_{n-2} and τ_{n-4} with τ_n links 6 cycles into one, breaking the parity barrier. We start with the 2 blue and 4 red thin cycles, and replace the dotted edges with the thick black solid edges to obtain the solid cycle. *Right:* the permutations appearing at the vertices marked with solid discs, listed in clockwise order starting and ending at the circled vertex, which is $[\ldots, a, b, c, d, e]$. The ellipsis $\cdots\cdot$ represents the same sequence everywhere it occurs.

cycles. Less obviously, it turns out that this parity is preserved by such a modification even if we relax the assumption that the q deleted edges lie in distinct cycles. (See [21] or [22] for proofs.) This is a problem, because many cycle covers that one might naturally start with have an *even* number of cycles. This holds in particular for the cycle cover that uses a single generator τ_k everywhere (for $n \geq 5$), and also for the one that arises in an obvious inductive approach to proving Theorem 1 (comprising $|A_n|/|A_{n-2}| = n(n-1)$ cycles each of length $|A_{n-2}|$). Thus we can (apparently) never get to a Hamiltonian cycle (i.e. a cycle cover of one cycle) by this method.

The above ideas in fact lead to the following rigorous condition for non-existence of directed Hamiltonian cycles. The result was proved by Rankin [21], based on an 1886 proof by Thompson [18] of a special case arising in change ringing; Swan [22] later gave a simpler version of the proof.

Theorem 2. *(Rankin) Consider the directed Cayley graph $\mathcal{G}$ of a finite group with two generators a, b. If $\operatorname{order}(ab^{-1})$ is odd and $|\mathcal{G}|/\operatorname{order}(a)$ is even, then $\mathcal{G}$ has no directed Hamiltonian cycle.*

An immediate consequence is that $\mathcal{A}_5$ has no directed Hamiltonian cycle (confirming the computer search result of [9]), and indeed $\mathcal{A}_n$ has no directed Hamiltonian cycle using only two generators for odd $n \geq 5$.

To break the parity barrier, we must use at least three generators in a fundamental way. The problem with the previous approach was that $\operatorname{order}(\tau_\ell \tau_k^{-1})$ is odd: we need an analogous relation involving composition of an *even* number of ratios of two generators. In terms of the graph $\mathcal{A}_n$, we need a cycle of length a multiple of 4 whose edges are oriented in alternating directions. It is clear that such a thing must exist for all odd $n \geq 7$, because the ratios $\tau_k \tau_\ell^{-1}$ generate the alternating group on the $n-2$ elements $\{3, \ldots, n\}$, which contains elements of even order. We will use the example:

$$\operatorname{order}(\zeta) = 2, \quad \text{where } \zeta := \tau_n \tau_{n-2}^{-1} \tau_n \tau_{n-4}^{-1} \tau_n \tau_{n-4}^{-1}. \tag{1}$$

It is a routine matter to check (1): the ratio $\tau_n \tau_{n-s}^{-1}$ is the permutation that jumps an element from position n to $n-s$ (while fixing $1, \ldots, n-s-1$ and shifting $n-s, \ldots, n-1$ right one place), so to compute the composition ζ of three such ratios we need only keep track of the last 5 elements. Fig. 2 (right) shows the explicit computation: starting from an arbitrary permutation $\pi = [\ldots, a, b, c, d, e] \in A_n$, the successive compositions $\pi, \pi\tau_n, \pi\tau_n\tau_{n-2}^{-1}, \pi\tau_n\tau_{n-2}^{-1}\tau_n, \ldots, \pi\zeta^2 = \pi$ are listed – the ellipsis $\cdots\cdot$ represents the same sequence everywhere it occurs. This explicit listing of the relevant permutations will be useful later.

We can use the above observation to link 6 cycles into one, as shown in Fig. 2 (left). Let Q' be a length-12 cycle in $\mathcal{A}_n$ with edges in alternating orientations that corresponds to the identity (1). That is to say, every alternate edge in Q' has label τ_n, and is oriented in the same direction around Q'. The other 6 edges are oriented in the opposite direction, and have successive labels $\tau_{n-2}, \tau_{n-4}, \tau_{n-4}, \tau_{n-2}, \tau_{n-4}, \tau_{n-4}$. Suppose that we start with a cycle cover in which the two τ_{n-2}-edges and the four τ_{n-4}-edges of Q' all lie in distinct cycles. Then we can delete these 6 edges and replace them with the six τ_n-edges of Q'. This results in a new cycle cover in which these 6 cycles have been united into one, thus reducing the number of cycles by 5 and changing its parity. See Fig. 2 (left) – the old cycles are in thin red and blue, while the new cycle is shown by solid lines and arcs.

We will prove Theorem 1 by induction. The inductive step will use one instance of the above 6-fold linkage to break the parity barrier, together with many instances of the simpler 3-fold linkage described earlier with $(k, \ell) = (n-2, n)$. The base case $n = 7$ will use the 6-fold linkage in the reverse direction (replacing six τ_n-edges with $\tau_{n-2}, \tau_{n-4}, \ldots$), together with the cases $(k, \ell) = (7, 5), (7, 3)$ of the earlier linkage.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2016.2620160, IEEE Transactions on Information Theory

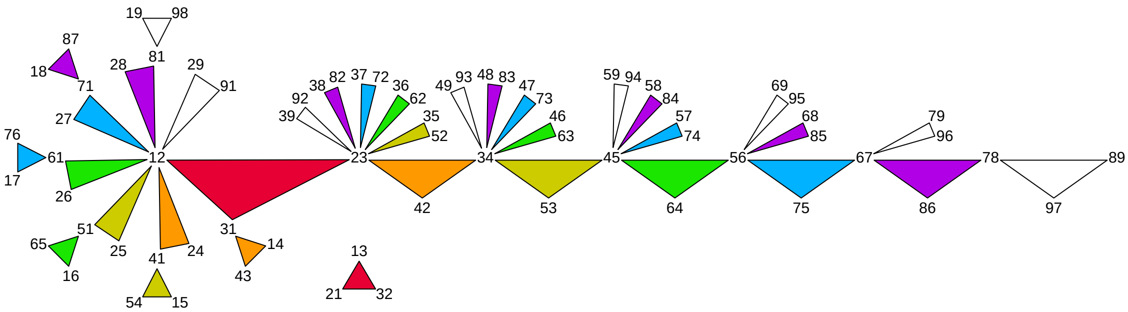

Fig. 3. The hypergraph of Proposition 3, when $n = 9$. The vertices are all the ordered pairs $(a, b) = ab \in [n]^{(2)}$, and the hyperedges are triangles of the form $\{ab, bc, ca\}$. Hyperedges are colored according to the step of the induction at which they are added. In the last step from $n = 8$ to $n = 9$, all the white hyperedges are added, i.e. those incident to vertices that contain element 9.

III. Hypergraph spanning

The other main ingredient for our proof is a systematic way of organizing the various linkages. For this the language of hypergraphs will be convenient. Similar hypergraph constructions were used in [9], [10]. A **hypergraph** (V, H) consists of a vertex set V and a set H of nonempty subsets of V, which are called **hyperedges**. A hyperedge of size r is called an r-hyperedge.

The **incidence graph** of a hypergraph (V, H) is the bipartite graph with vertex set $V \cup H$, and with an edge between $v \in V$ and $h \in H$ if $v \in h$. A **component** of a hypergraph is a component of its incidence graph, and a hypergraph is **connected** if it has one component. We say that a hypergraph is **acyclic** if its incidence graph is acyclic. Note that this a rather strong condition: for example, if two distinct hyperedges h and h' share two distinct vertices v and v' then the hypergraph is not acyclic. (Several non-equivalent notions of acyclicity for hypergraphs have been considered – the notion we use here is sometimes called Berge-acyclicity – see e.g. [26]).

We are interested in hypergraphs of a particular kind that are related to the linkages considered in the previous section. Let $[n]^{(k)}$ be the set of all $n!/(n-k)!$ ordered k-tuples of distinct elements of $[n]$. If $t = (a, b, c) \in [n]^{(3)}$ is a triple, define the **triangle** $\Delta(t) = \Delta(a, b, c) := \{(a, b), (b, c), (c, a)\} \subset [n]^{(2)}$ of pairs that respect the cyclic order. (Note that $\Delta(a, b, c) = \Delta(c, a, b) \neq \Delta(c, b, a)$.) In our application to Hamiltonian cycles, $\Delta(a, b, c)$ will encode precisely the linkage of 3 cycles shown in Fig. 1. The following fact and its proof are illustrated in Fig. 3.

Proposition 3. *Let $n \geq 3$. There exists an acyclic hypergraph with vertex set $[n]^{(2)}$, with all hyperedges being triangles $\Delta(t)$ for $t \in [n]^{(3)}$, and with exactly two components: one containing precisely the 3 vertices of $\Delta(3, 2, 1)$, and the other containing all other vertices.*

Proof. We give an explicit inductive construction. When $n = 3$ we simply take as hyperedges the two triangles $\Delta(3, 2, 1)$ and $\Delta(1, 2, 3)$.

Now let $n \geq 4$, and assume that $([n-1]^{(2)}, H)$ is a hypergraph satisfying the given conditions for $n - 1$. Consider the larger hypergraph $([n]^{(2)}, H)$ with the same set of hyperedges, and note that its components are precisely: (i) $\Delta(3, 2, 1)$; (ii) an acyclic component which we denote K that contains all vertices of $[n-1]^{(2)} \setminus \Delta(3, 2, 1)$; and (iii) the $2n - 2$ isolated vertices $\{(i, n), (n, i) : i \in [n-1]\}$.

We will add some further hyperedges to $([n]^{(2)}, H)$. For $i \in [n-1]$, write i^+ for the integer in $[n-1]$ that satisfies $i^+ \equiv (i + 1) \bmod (n - 1)$, and define

$$\begin{aligned} D &:= \{\Delta(i, i^+, n) : i \in [n-1]\} \\ &= \{\Delta(1, 2, n), \Delta(2, 3, n), \ldots \\ &\qquad \ldots, \Delta(n-2, n-1, n),\ \Delta(n-1, 1, n)\}. \end{aligned}$$

Any element $\Delta(i, i^+, n)$ of D has 3 vertices. One of them, (i, i^+), lies in K, while the others, (i^+, n) and (n, i), are isolated in $([n]^{(2)}, H)$. Moreover, each isolated vertex of $([n]^{(2)}, H)$ appears in exactly one hyperedge in D. Therefore, $([n]^{(2)}, H \cup D)$ has all the claimed properties. □

We remark that the above hypergraph admits a simple (non-inductive) description – it consists of all $\Delta(a, b, c)$ such that $\max\{a, b\} < c$ and $b \equiv (a + 1) \bmod (c - 1)$.

In order to link cycles into a Hamiltonian cycle we will require a *connected* hypergraph. For $n \geq 3$ there is no connected acyclic hypergraph of triangles with vertex set $[n]^{(2)}$. (This follows from parity considerations: an acyclic component composed of m triangles has $1 + 2m$ vertices, but $|[n]^{(2)}|$ is even.) Instead, we simply introduce a larger hyperedge, as follows.

Corollary 4. *Let $n \geq 5$ and let $a, b, c, d, e \in [n]$ be distinct. There exists a connected acyclic hypergraph with vertex set $[n]^{(2)}$ such that one hyperedge is the 6-hyperedge $\Delta(a, b, e) \cup \Delta(c, d, e)$, and all others are triangles $\Delta(t)$ for $t \in [n]^{(3)}$.*

Proof. By symmetry, it is enough to prove this for any one choice of (a, b, c, d, e); we choose $(2, 1, 4, 5, 3)$. The result follows from Proposition 3, on noting that $\Delta(3, 4, 5) = \Delta(4, 5, 3)$ is a hyperedge of the hypergraph constructed there: we simply unite it with $\Delta(3, 2, 1) = \Delta(2, 1, 3)$ to form the 6-hyperedge. □

IV. The Hamiltonian cycle

We now prove Theorem 1 by induction on (odd) n. We give the inductive step first, followed by the base case $n = 7$. The following simple observation will be used in the inductive step.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2016.2620160, IEEE Transactions on Information Theory

Lemma 5. *Let $n \geq 3$ be odd, and consider any Hamiltonian cycle of $\mathcal{A}_n$. For every $i \in [n]$ there exists a permutation $\pi \in A_n$ with $\pi(n) = i$ that is immediately followed by a τ_n-edge in the cycle.*

Proof. Since the cycle visits all permutations of A_n, it must contain a directed edge from a permutation π satisfying $\pi(n) = i$ to a permutation π' satisfying $\pi'(n) \neq i$. This is a τ_n-edge, since any other generator would fix the rightmost element. □

Proof of Theorem 1, inductive step. We will prove by induction on odd $n \geq 7$ the statement:

$$\begin{array}{c}\textit{there exists a Hamiltonian cycle of } \mathcal{A}_n \\ \textit{that includes at least one } \tau_{n-2}\textit{-edge.}\end{array} \tag{2}$$

As mentioned above, we postpone the proof of the base case $n = 7$. For distinct $a, b \in [n]$ define the set of permutations of the form $[\dots, a, b]$:

$$A_n(a, b) := \Big\{\pi \in A_n : \big(\pi(n-1), \pi(n)\big) = (a, b)\Big\}.$$

Let $n \geq 9$, and let $L = (\tau_{s(1)}, \tau_{s(2)}, \dots, \tau_{s(m)})$ be the sequence of generators used by a Hamiltonian cycle of $\mathcal{A}_{n-2}$, as guaranteed by the inductive hypothesis, in the order that they are encountered in the cycle starting from $\mathrm{id} \in A_{n-2}$ (where $m = (n-2)!/2$, and $s(i) \in \{3, 5, \dots, n-2\}$ for each i). Now start from any permutation $\pi \in A_n(a, b)$ and apply the sequence of generators L (where a generator $\tau_k \in A_{n-2}$ is now interpreted as the generator $\tau_k \in A_n$ with the same name). This gives a cycle in $\mathcal{A}_n$ whose vertex set is precisely $A_n(a, b)$. (The two rightmost elements a, b of the permutation are undisturbed, because L does not contain τ_n.) Note that, for given a, b, different choices of the starting vertex π in general result in different cycles.

We next describe the idea of the proof, before giving the details. Consider a cycle cover $\mathcal{C}$ comprising, for each $(a, b) \in [n]^{(2)}$, one cycle $C(a, b)$ with vertex set $A_n(a, b)$ of the form described above (so $n(n-1)$ cycles in total). We will link the cycles of $\mathcal{C}$ together into a single cycle by substituting the generator τ_n at appropriate points, in the ways discussed in Section II. The linking procedure will be encoded by the hypergraph of Corollary 4. The vertex (a, b) of the hypergraph will correspond to the initial cycle $C(a, b)$. A 3-hyperedge $\Delta(a, b, c)$ will indicate a substitution of τ_n for τ_{n-2} in 3 of the cycles of $\mathcal{C}$, linking them together in the manner of Fig. 1. The 6-hyperedge will correspond to the parity-breaking linkage in which τ_n is substituted for occurrences of both τ_{n-2} and τ_{n-4}, linking 6 cycles as in Fig. 2. One complication is that the starting points of the cycles of $\mathcal{C}$ must be chosen so that τ_{n-2}- and τ_{n-4}-edges occur in appropriate places so that all these substitutions are possible. To address this, rather than choosing the cycle cover $\mathcal{C}$ at the start, we will in fact build our final cycle sequentially, using one hyperedge at a time, and choosing appropriate cycles $C(a, b)$ as we go. We will start with the 6-hyperedge, and for each subsequent 3-hyperedge we will link in two new cycles. Lemma 5 will ensure enough τ_{n-2}-edges for subsequent steps: for any $(a, b, c) \in [n]^{(3)}$, there is a vertex of the form $[\dots, a, b, c]$ in $C(b, c)$ followed by a τ_{n-2}-edge. The inductive hypothesis (2) will provide the τ_{n-4}-edges needed for the initial 6-fold linkage.

We now give the details. In preparation for the sequential linking procedure, choose an acyclic connected hypergraph $([n]^{(2)}, H)$ according to Corollary 4, with the 6-hyperedge being $\Delta_0 \cup \Delta_0'$, where $\Delta_0 := \Delta(c, d, e)$ and $\Delta_0' := \Delta(a, b, e)$, and where we write

$$(a, b, c, d, e) = (n-4, n-3, n-2, n-1, n). \tag{3}$$

Let $N = |H| - 1$, and order the hyperedges as $H = \{h_0, h_1, \dots, h_N\}$ in such a way that $h_0 = \Delta_0 \cup \Delta_0'$ is the 6-hyperedge, and, for each $1 \leq i \leq N$, the hyperedge h_i shares exactly one vertex with $\bigcup_{\ell=0}^{i-1} h_\ell$. (To see that this is possible, note that for any choice of $h_0, \dots, h_{i-1}$ satisfying this condition, connectedness of the hypergraph implies that there exists h_i that shares *at least* one vertex with one of its predecessors; acyclicity then implies that it shares exactly one.)

We will construct the required Hamiltonian cycle via a sequence of steps $j = 0, \dots, N$. At the end of step j we will have a self-avoiding directed cycle C_j in $\mathcal{A}_n$ with the following properties.

(i) The vertex set of C_j is the union of $A_n(x, y)$ over all $(x, y) \in \bigcup_{i=0}^{j} h_i$.

(ii) For every $(x, y, z) \in [n]^{(3)}$ such that $(y, z) \in \bigcup_{i=0}^{j} h_i$ but $\Delta(x, y, z) \notin \{\Delta_0, \Delta_0', h_1, h_2, \dots, h_j\}$, there exists a permutation $\pi \in A_n$ of the form $[\dots, x, y, z]$ that is followed immediately by a τ_{n-2}-edge in C_j.

We will check by induction on j that the above properties hold. The final cycle C_N will be the required Hamiltonian cycle. The purpose of the technical condition (ii) is to ensure that suitable edges are available for later linkages; the idea is that the triple (x, y, z) is available for linking in two further cycles unless it has already been used.

We will describe the cycles C_j by giving their sequences of generators. Recall that L is the sequence of generators of the Hamiltonian cycle of $\mathcal{A}_{n-2}$. Note that L contains both τ_{n-2} and τ_{n-4}, by Lemma 5 and the inductive hypothesis (2) respectively. For each of $k = n-2, n-4$, fix some location i where τ_k occurs in L (so that $s(i) = k$), and let $L[\tau_k]$ be the sequence obtained by starting at that location and omitting this τ_k from the cycle:

$$L[\tau_k] := \big(\tau_{s(j+1)}, \tau_{s(j+2)} \dots, \tau_{s(m)},\ \tau_{s(1)}, \dots, \tau_{s(j-1)}\big).$$

Note that the composition in order of the elements of $L[\tau_k]$ is τ_k^{-1}.

For step 0, let C_0 be the cycle that starts at $\mathrm{id} \in A_n$ and uses the sequence of generators

$$\begin{array}{c}\tau_n, L[\tau_{n-2}], \tau_n, L[\tau_{n-4}], \tau_n, L[\tau_{n-4}], \\ \tau_n, L[\tau_{n-2}], \tau_n, L[\tau_{n-4}], \tau_n, L[\tau_{n-4}],\end{array}$$

(where commas denote concatenation). This cycle is precisely of the form illustrated in Fig. 2 (left) by the solid arcs and lines. The curved arcs represent the paths corresponding to the $L[\cdot]$ sequences. The vertex set of each such path is precisely $A_n(u, v)$ for some pair (u, v); we denote this

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2016.2620160, IEEE Transactions on Information Theory

path $P(u,v)$. The solid lines represent the τ_n-edges. Moreover, since Fig. 2 (right) lists the vertices (permutations) at the beginning and end of each path $P(u,v)$, we can read off the pairs (u,v). With $a,\dots,e$ as in (3), the pairs are $\{(d,e),(c,d),(e,c),(b,e),(a,b),(e,a)\}$. This set equals $\Delta_0 \cup \Delta_0' = h_0$, so property (i) above holds for the cycle C_0.

We next check that C_0 satisfies (ii). Let $(x,y,z) \in [n]^{(3)}$ be such that $(y,z) \in h_0$. The cycle C_0 includes a path $P(y,z)$ with vertex set $A_n(y,z)$ and generator sequence $L[\tau_k]$ (where k is $n-2$ or $n-4$). Let $C(y,z)$ be the cycle that results from closing the gap, i.e. appending a τ_k-edge f to the end of $P(y,z)$. Note that $P(y,z)$ and $C(y,z)$ both have vertex set $A_n(y,z)$. By Lemma 5 applied to $\mathcal{A}_{n-2}$, the cycle $C(y,z)$ contains a permutation of the form $[\dots,x,y,z]$ immediately followed by a τ_{n-2}-edge, g say. Edge g is also present in C_0 unless $g=f$. Consulting Fig. 2, and again using the notation in (3), we see that this happens only in the two cases $(x,y,z)=(e,c,d),(e,a,b)$. But in these cases we have $\Delta(x,y,z)=T_0,T_0'$ respectively. Thus condition (ii) is satisfied at step 0.

Now we inductively describe the subsequent steps. Suppose that step $j-1$ has been completed, giving a cycle C_{j-1} that satisfies (i) and (ii) (with parameter $j-1$ in place of j). We will augment C_{j-1} to obtain a larger cycle C_j, in a manner encoded by the hyperedge h_j. Let

$$h_j = \Delta(a,b,c) = \big\{(a,b),(b,c),(c,a)\big\}$$

(where we no longer adopt the notation (3)). By our choice of the ordering of H, exactly one of these pairs belongs to $\bigcup_{i=0}^{j-1} h_i$; without loss of generality, let it be (b,c). By property (ii) of the cycle C_{j-1}, it contains a vertex of the form $[\dots,a,b,c]$ immediately followed by a τ_{n-2}-edge, f say. Delete edge f from C_{j-1} to obtain a directed path P_{j-1} with the same vertex set. Append to P_{j-1} the directed path that starts at the endvertex of P_{j-1} and then uses the sequence of generators

$$\tau_n, L[\tau_{n-2}], \tau_n, L[\tau_{n-2}], \tau_n.$$

Since $\operatorname{order}(\tau_n\tau_{n-2}^{-1})=3$, this gives a cycle, which we denote C_j.

The new cycle C_j has precisely the form shown in Fig. 1 (left) by the solid arcs and lines, where C_{j-1} is the thin blue cycle in the upper left, containing the circled vertex, which is the permutation $[\dots,a,b,c]$. The arc is P_{j-1}, and the dotted edge is f. As before, the permutations at the filled discs may be read from Fig. 1 (right). Thus, C_j consists of the path P_{j-1}, together with two paths $P(a,b),P(c,a)$ with respective vertex sets $A_n(a,b),A_n(c,a)$ (the other two thin blue arcs in the figure), and three τ_n-edges (thick black lines) connecting these three paths. Hence C_j satisfies property (i).

We now check that C_j satisfies (ii). The argument is similar to that used in step 0. Let (x,y,z) satisfy the assumptions in (ii). We consider two cases. First suppose $(y,z) \in \bigcup_{i=0}^{j-1} h_i$. Then property (ii) of C_{j-1} implies that C_{j-1} has a vertex of the form $[\dots,x,y,z]$ followed by a τ_{n-2}-edge g, say. Then g is also present in C_j unless $g=f$. But in that case we have $(x,y,z)=(a,b,c)$, and so $\Delta(x,y,z)=h_j$, contradicting the assumption on (x,y,z). On the other hand, suppose $(y,z) \in h_j \setminus \bigcup_{i=0}^{j-1} h_i$. Then (y,z) equals (a,b) or (c,a). Suppose the former; the argument in the latter case is identical. Let $C(a,b)$ be the cycle obtained by appending a τ_{n-2}-edge to $P(a,b)$. Applying Lemma 5 shows that $C(a,b)$ contains a vertex of the form $[\dots,x,a,b]$ followed by a τ_{n-2}-edge g, say. Then g is also present in $P(a,b)$ unless $x=c$, but in that case $\Delta(x,y,z)=h_j$, contradicting the assumption in (ii). Thus, property (ii) is established.

To conclude the proof, note that the final cycle C_N is Hamiltonian, by property (i) and the fact that the hypergraph of Corollary 4 has vertex set $[n]^{(2)}$. To check that it includes some τ_{n-2}-edge as required for (2), recall that h_N has only one vertex in common with $h_0,\dots,h_{N-1}$, so there exist x,y,z with $(y,z) \in h_N$ but $\Delta(x,y,z) \notin H$. Hence property (ii) implies that C_N contains a τ_{n-2}-edge. □

TABLE I
RULES FOR GENERATING A DIRECTED HAMILTONIAN CYCLE OF $\mathcal{A}_7$.

row	permutations	generator
1	$6\overline{777}{*}{*}{*}, \overline{777}6{*}{*}{*}$	τ_5
2	$67{*}{*}{*}{*}{*}, 76{*}{*}{*}{*}{*}$	τ_3
3	$567\overline{1}{*}{*}{*}, 576{*}{*}{*}{*}$	τ_5
4	$2567{*}{*}{*}, 4576{*}{*}{*}$	τ_5
5	$5671234, 5612347, 5623714, 5637142$	τ_3
6	$5623471, 5671423$	τ_5
7	otherwise	τ_7

Permutations of the given forms should be followed by the generator in the same row of the table. The symbol $*$ denotes an arbitrary element of [7], and $\overline{a}$ denotes any element other than a.

Proof of Theorem 1, base case. For the base case of the induction, we give an explicit directed Hamiltonian cycle of $\mathcal{A}_7$ that includes τ_5 at least once. (In fact the latter condition must necessarily be satisfied, since, as remarked earlier, Theorem 2 implies that there is no Hamiltonian cycle using only τ_3 and τ_7.)

Table I specifies which generator the cycle uses immediately after each permutation of A_7, as a function of the permutation itself. The skeptical reader may simply check by computer that these rules generate the required cycle. But the rules were constructed by hand; below we briefly explain how.

First suppose that from every permutation of A_7 we use the outgoing τ_7-edge, as specified in row 7 of the table. This gives a cycle cover comprising $|A_7|/7 = 360$ cycles of size 7. Now consider the effect of replacing some of these τ_7's according to rows 1–6 in succession. Each such replacement performs a linkage, as discussed in Section II. Row 1 links the cycles in sets of 3 to produce 120 cycles of length 21, each containing exactly one permutation of the form $67{*}{*}{*}{*}{*}$ or $76{*}{*}{*}{*}{*}$. Row 2 then links these cycles in sets of 5 into 24 cycles of length 105, each containing exactly one permutation of the form $675{*}{*}{*}{*}$ or $765{*}{*}{*}{*}$. Rows 3 and 4 link various sets of three cycles, permuting elements 1234, to produce 6 cycles. Finally, rows 5 and 6 break the parity barrier as discussed earlier, uniting these 6 cycles into one. □

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2016.2620160, IEEE Transactions on Information Theory

TABLE II

A CYCLE OF LENGTH 315 IN $\mathcal{S}_6$ WITH NO TWO PERMUTATIONS RELATED BY AN ELEMENTARY TRANSPOSITION.

$$\big(64\ 5^2353^25^435^335^33^25^435^435^335^43^25^335^43^2$$
$$64\ 5^33^2535^23^25^23^2535^335^435^335^33^25^435^335^43^25\big)^3.$$

The cycle uses the sequence of generators $(\tau_{k(i)})$ where $(k(i))_{i=1}^{315}$ is the given sequence. Commas are omitted, and superscripts indicate repetitions.

V. EVEN SIZE

We briefly discuss a possible approach for even n. Recall that M_n is the maximum length of a cycle $\mathcal{S}_n$ in which no two permutations are related by an adjacent transposition.

To get a cycle longer than M_{n-1} we must use τ_n. But this is an odd permutation, so we cannot remain in the alternating group A_n. We suggest following τ_n immediately by another odd generator, say τ_{n-2}, in order to return to A_n (note that τ_2 is forbidden). In order to include permutations of the form $[\ldots, j]$ for every $j \in [n]$, we need to perform such a transition (at least) n times in total in our cycle. In the ith transition we visit one odd permutation, α_i say, between the generators τ_n and τ_{n-2}. For the remainder of the cycle we propose using only generators τ_k for odd k, so that we remain in A_n.

One may fix the permutations $\alpha_1, \ldots, \alpha_n$ in advance. The problem then reduces to that of finding long self-avoiding directed paths in $\mathcal{A}_{n-1}$, with specified start and end vertices, and avoiding certain vertices – those that would result in a permutation that is related to some α_i by an elementary transposition. There are $O(n^2)$ vertices to be avoided in total.

Since, for large n, the number of vertices to be avoided is much smaller than $|A_{n-1}|$, we think it very likely that paths of length $(1 - o(1))|A_{n-1}|$ exist, which would give $M_n \geq (1 - o(1))n!/2$ as $n \to \infty$. It is even plausible that $M_n \geq n!/2 - O(n^2)$ might be achievable. The graph $\mathcal{A}_{n-1}$ seems to have a high degree of global connectivity, as evidenced by the diverse constructions of cycles of close to optimal length in [8]–[10]. For a specific approach (perhaps among others), one might start with a short path linking the required start and end vertices, and then try to successively link in short cycles (say those that use a single generator such as τ_{n-1}) in the manner of Fig. 1, leaving out the relatively few short cycles that contain forbidden vertices. It is conceivable that the forbidden vertices might conspire to prevent this, for example by blocking even short paths between the start and end vertices. However, this appears unlikely, especially given the flexibility in the choice of $\alpha_1, \ldots, \alpha_n$.

While there appear to be no fundamental obstacles, a proof for general even n along the above lines would likely be rather messy. (Of course, this does not preclude some other approach). Instead, the approach was combined with a computer search to obtain a cycle of length $315 = 6!/2 - 45$ for $n = 6$, given in Table II, answering a question of [9], and improving the previous record $M_6 \geq 57$ [9] by more than a factor of 5. The case $n = 6$ is in some respects harder than larger n: the forbidden vertices form a larger fraction of the total, and $\mathcal{A}_5$ has only two generators, reducing available choices. (On the other hand, the search space is of course relatively small). Thus, this result also lends support to the belief that $M_n \geq (1 - o(1))n!/2$ as $n \to \infty$.

The search space was reduced by quotienting the graph $\mathcal{S}_6$ by a group of order 3 to obtain a Schreier graph, giving a cycle in which the sequence of generators is repeated 3 times. The cycle is given in Table II.

REFERENCES

[1] A. Jiang, R. Mateescu, M. Schwartz, and J. Bruck, "Rank modulation for flash memories," *IEEE Trans. Inform. Theory*, vol. 55, no. 6, pp. 2659–2673, 2009.

[2] A. E. Holroyd, F. Ruskey, and A. Williams, "Shorthand universal cycles for permutations," *Algorithmica*, vol. 64, no. 2, pp. 215–245, 2012.

[3] D. E. Knuth, *The art of computer programming. Vol. 4, Fasc. 2, Generating all tuples and permutations.* Addison-Wesley, 2005.

[4] A. Mazumdar, A. Barg, and G. Zémor, "Constructions of rank modulation codes," *IEEE Trans. Inform. Theory*, vol. 59, no. 2, pp. 1018–1029, 2013.

[5] C. Savage, "A survey of combinatorial Gray codes," *SIAM Rev.*, vol. 39, no. 4, pp. 605–629, 1997.

[6] J. Baylis, *Error-correcting codes, a mathematical introduction.* Chapman & Hall, London, 1998.

[7] H. L. Abbott and M. Katchalski, "On the snake in the box problem," *J. Combin. Theory Ser. B*, vol. 45, no. 1, pp. 13–24, 1988.

[8] Y. Yehezkeally and M. Schwartz, "Snake-in-the-box codes for rank modulation," *IEEE Trans. Inform. Theory*, vol. 58, no. 8, pp. 5471–5483, 2012.

[9] M. Horovitz and T. Etzion, "Constructions of snake-in-the-box codes for rank modulation," *IEEE Trans. Inform. Theory*, vol. 60, no. 11, pp. 7016–7025, 2014.

[10] Y. Zhang and G. Ge, "Snake-in-the-box codes for rank modulation under Kendall's τ-metric," 2015, arXiv:1506.02740.

[11] M. G. Kendall, "A new measure of rank correlation," *Biometrika*, vol. 30, no. 1/2, pp. 81–93, 1938.

[12] A. Williams, "Hamiltonicity of the Cayley digraph on the symmetric group generated by $\sigma = (12 \cdots n)$ and $\tau = (12)$," 2013, arXiv:1307.2549.

[13] F. Ruskey and A. Williams, "The coolest way to generate combinations," *Discrete Mathematics*, vol. 309, no. 17, pp. 5305–5320, 2009.

[14] I. Pak and R. Radoičić, "Hamiltonian paths in Cayley graphs," *Discrete Mathematics*, vol. 309, no. 17, pp. 5501–5508, 2009.

[15] S. J. Curran and J. A. Gallian, "Hamiltonian cycles and paths in Cayley graphs and digraphs – a survey," *Discrete Mathematics*, vol. 156, no. 1, pp. 1–18, 1996.

[16] D. Witte and J. A. Gallian, "A survey: Hamiltonian cycles in Cayley graphs," *Discrete Math.*, vol. 51, no. 3, pp. 293–304, 1984.

[17] A. T. White, "Fabian Stedman: the first group theorist?" *Amer. Math. Monthly*, vol. 103, no. 9, pp. 771–778, 1996.

[18] W. H. Thompson, "A note on Grandsire Triples," Macmillan and Bowes, Cambridge, 1886, reprinted in W. Snowdon, *Grandsire*, London 1905.

[19] D. Griffiths, "Twin bob plan composition of Stedman Triples: Partitioning of graphs into Hamiltonian subgraphs," Research Report 94(37), Univ. of Sydney, School of Math. and Stat., 1994. [Online]. Available: www.maths.usyd.edu.au/res/CompAlg/Gri/2bob-sted.html

[20] R. Duckworth and F. Stedman, *Tintinnalogia, or, the art of ringing*, London, 1667. [Online]. Available: www.gutenberg.org/ebooks/18567

[21] R. A. Rankin, "A campanological problem in group theory," *Proc. Cambridge Philos. Soc.*, vol. 44, pp. 17–25, 1948.

[22] R. G. Swan, "A simple proof of Rankin's campanological theorem," *Amer. Math. Monthly*, vol. 106, no. 2, pp. 159–161, 1999.

[23] E. Rapaport-Strasser, "Cayley color groups and Hamilton lines," *Scripta Math.*, vol. 24, pp. 51–58, 1959.

[24] R. C. Compton and S. G. Williamson, "Doubly adjacent Gray codes for the symmetric group," *Linear and Multilinear Algebra*, vol. 35, no. 3-4, pp. 237–293, 1993.

[25] T. Mütze, "Proof of the middle levels conjecture," *Proc. Lond. Math. Soc. (3)*, vol. 112, no. 4, pp. 677–713, 2016.

[26] R. Fagin, "Degrees of acyclicity for hypergraphs and relational database schemes," *J. Assoc. Comput. Mach.*, vol. 30, no. 3, pp. 514–550, 1983.

Alexander E. Holroyd received his PhD from the University of Cambridge in 2000. Prior to joining the Theory Group of Microsoft Research, he held positions at the University of California in Los Angeles and Berkeley and the University of British Columbia.