Discrete Applied Mathematics 159 (2011) 1641–1645

Contents lists available at ScienceDirect

Discrete Applied Mathematics

journal homepage: www.elsevier.com/locate/dam

Note

Girth of pancake graphs

Phillip E.C. Compeau *

Department of Mathematics, Davidson College, Davidson, NC 28035, USA

ARTICLE INFO

Article history:
Received 1 March 2010
Received in revised form 20 March 2011
Accepted 13 June 2011
Available online 13 July 2011

Keywords:
Pancake graphs
Girth
Reversals
Breakpoints

ABSTRACT

We consider four families of pancake graphs, which are Cayley graphs, whose vertex sets are either the symmetric group on n objects or the hyperoctahedral group on n objects and whose generating sets are either all reversals or all reversals inverting the first k elements (called prefix reversals). We find that the girth of each family of pancake graphs remains constant after some small threshold value of n.

1. Introduction

A *signed permutation on n objects* is a function

$$\alpha: \{1, 2, \ldots, n\} \to \{-n, \ldots, -1\} \cup \{1, \ldots, n\}$$

such that $|\alpha|$ is in S_n. We represent a signed permutation α as an n-tuple

$$\alpha = (\alpha(1), \alpha(2), \ldots, \alpha(n))$$

and we can think about α as a permutation on n objects in which each object is provided a sign. For two signed permutations on n objects, say α and β, we define the composition of β with α by

$$(\beta\alpha)(i) := \alpha(|\beta(i)|) \cdot \mathrm{sgn}\beta(i).$$

For example, $(-4, 1, -3, -2)(1, 3, -4, -2) = (2, 1, 4, -3)$. Under this operation, the set of all signed permutations on n objects forms the *hyperoctahedral group on n objects*, which we denote by B_n. This group is commonly known as the group of symmetries of the n-dimensional hypercube.

In this light, we shall call the members of S_n *unsigned permutations on n objects*. As with composition of signed permutations, our composition in S_n will be written left-to-right:

$$(\beta\alpha)(i) := \alpha(\beta(i)).$$

Henceforth, "permutation" will refer generally to both signed and unsigned permutations. The identity of both S_n and B_n is denoted by I_n.

For fixed n and $1 \le j < k \le n$, the *unsigned reversal on the interval* $[j, k]$ is the permutation $\upsilon_{[j,k]} \in S_n$ defined by

$$\upsilon_{[j,k]} := (1, 2, \ldots, j-1, k, k-1, \ldots, j, k+1, k+2, \ldots, n).$$

* Corresponding address: UC-San Diego, Department of Mathematics, 9500 Gilman Drive #0112, La Jolla, CA 92093, USA. Tel.: +1 336 984 1266; fax: +1 858 534 5273.

E-mail address: pcompeau@math.ucsd.edu.

0166-218X/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.dam.2011.06.013

In the case that $j = 1$ (and thus $2 \le k \le n$), we write $\upsilon_{[k]} := \upsilon_{[1,k]}$ and say that $\upsilon_{[k]}$ is the *unsigned prefix reversal at index k*. Let Υ_n (Υ_n^p) denote the set of all unsigned reversals (unsigned prefix reversals) in S_n.

Analogously, for $1 \le j \le k \le n$, the *signed reversal on the interval* $[j, k]$ is the permutation $\sigma_{[j,k]} \in B_n$ defined by

$$\sigma_{[j,k]} := (1, 2, \ldots, j-1, -k, -(k-1), \ldots, -j, k+1, k+2, \ldots, n).$$

Note that in the signed case we allow $k = j$, since for all $1 \le j \le n$, $\sigma_{[j,j]} \ne I_n$. Furthermore, for $1 \le k \le n$, the *signed prefix reversal at index k* is the permutation $\sigma_{[k]} := \sigma_{[1,k]}$. Let Σ_n (Σ_n^p) denote the set of all signed reversals (signed prefix reversals) in B_n. Henceforth, "reversal" may refer to any signed or unsigned reversal, prefix or otherwise, and we use ρ to denote an arbitrary reversal.

Next, define the *unsigned reversal graph* (*unsigned prefix reversal graph*) on n objects, denoted by UR_n (UP_n), as the Cayley graph whose vertex set is S_n and whose generating set is Υ_n (Υ_n^p). Both graphs have order $|S_n| = n!$; UP_n has degree $n-1$, while UR_n has degree $\binom{n}{2}$.

Analogously, define the *signed reversal graph* (*signed prefix reversal graph*) on n objects, denoted by SR_n (SP_n), as the Cayley graph with vertex set B_n and generating set Σ_n (Σ_n^p). Both graphs have order $|B_n| = 2^n \cdot n!$; SP_n has degree n, while SR_n has degree $\binom{n+1}{2}$.

We refer to all four families of Cayley graphs collectively as *pancake graphs*. Before continuing, we will give two straightforward facts about pancake graphs. We will see later that the second fact will provide an upper bound on the girth of each family of pancake graphs once we have found a short cycle in that family.

Fact 1. *For any n, $\Upsilon_n^p \subset \Upsilon_n$ and $\Sigma_n^p \subset \Sigma_n$, so that UP_n is a subgraph of UR_n and SP_n is a subgraph of SR_n.*

Fact 2. *Every pancake graph embeds in all higher-order pancake graphs of the same family. For example, if $m \le n$, then UP_m is isomorphic to the Cayley subgraph of UP_n generated by the subset of Υ_n^p containing only those unsigned prefix reversals υ_k for which $k \le m$.*

The etymology of "pancake graph" traces back to a 1975 *American Mathematical Monthly* problem, which asked for a function $f(n)$ bounding the maximum number of flips required to transform a given stack of n differently sized pancakes into the stack whose pancakes are sorted from top to bottom in the order of increasing size. Of course, stacks of pancakes correspond to unsigned permutations; if all the pancakes are burned on one side (creating the "burned pancake problem"), stacks correspond to signed permutations. A pancake graph is therefore a graph whose vertices are stacks of n pancakes, and whose edges represent flips between stacks: prefix reversals constitute a "one-spatula" case and reversals constitute a "two-spatula" case. For three decades, the best known bound for the pancake problem was found in [6], although it has recently been improved in [3]. See [1,2,4,5,7,9] and [10] for more on the pancake problem and its offshoots.

A biological application of pancake flipping is found in genetic analysis. One common form of large-scale evolutionary change is a genomic mutation which manifests itself in the reversal of some segment of the mutated organism's DNA. Phylogeneticists study how the accumulation of millions of years of mutations, including reversals of DNA segments, have led to species divergence. Therefore, a given property of pancake graphs can in many cases be translated into a phylogenetic application when only reversal mutations are considered. For example, it seems extremely unlikely that evolutionary changes occur in cycles; however, if pancake graphs were to be shown to have large girth (say $O(n)$), then we would have concrete evidence that cyclical evolutionary patterns are implausible.

2. Preliminaries

In this note, our aim is to find the girth of each of the four families of pancake graphs introduced above. For $k \ge 3$, we define a *cycle of length k* in a pancake graph as a reduced finite sequence of reversals $(\rho_1, \rho_2, \ldots, \rho_k)$, all of which lie in the appropriate generating set, and such that $\rho_k \cdots \rho_1 = I_n$. By a "reduced sequence", we mean a sequence for which $\rho_1 \ne \rho_k$ and $\rho_{i+1} \ne \rho_i$ for all $1 \le i \le k-1$, since all reversals are involutions within their respective permutation groups. Observe that this definition of cycle agrees with the graph theoretical one; therefore, the girth of a pancake graph will be the minimal length of a cycle of reversals taken from that graph's generating set.

Let $\alpha \in S_n$ and extend α to a member of S_{n+2} by setting $\alpha(0) = 0$ and $\alpha(n+1) = n+1$. For $0 \le i \le n$, we continue the terminology established in [9] and say that α has a *breakpoint at i* if $|\alpha(i+1) - \alpha(i)| \ne 1$. In the signed case, we also extend $\beta \in B_n$ to an element of B_{n+2} by setting $\beta(0) = 0$ and $\beta(n+1) = n+1$. In this case, however, we say that β has a breakpoint at i if $\beta(i+1) - \beta(i) \ne 1$. For example, the signed permutation $\beta = (-4, -3, -2, 1, 5)$ has breakpoints at 0, 3, and 4.

We also define a *non-initial breakpoint* of a permutation to be any breakpoint other than the breakpoint at 0. Observe that in both the signed and unsigned cases, the identity I_n is the only permutation with no breakpoints. In pancake flipping, it is therefore interesting to think of breakpoints as something we wish to eliminate during a walk to the identity.

A set $B \subset B_n$ will be called k-compressible if for some $J \subset \{0, 1, 2, \ldots, n\}$ with $|J| = k$, B contains only signed permutations β such that for every $j \in J$, either $\beta^{-1}(j+1) - \beta^{-1}(j) = 1$ or $\beta^{-1}(-j) - \beta^{-1}(-(j+1)) = 1$. (In other words, if j occurs in β, then $j+1$ occurs immediately to the right of j in β, and otherwise $-(j+1)$ occurs immediately to the left of $-j$.) Note that at most one of these differences can be defined, and it may well be the case that neither is. To emphasize the importance of the choice of J, we may also call B J-compressible. For example, with $J = \{0, 2, 3\}$, the following subset of B_5 is J-compressible:

$$B = \{(1, -4, -3, -2, -5), (1, 2, 3, 4, -5), (1, 5, 2, 3, 4), (1, 2, 3, 4, 5), (1, -5, -4, -3, -2)\}.$$

In fact, for each J, there exists a unique maximum J-compressible set of B_n under the partial order of set inclusion. If $|J| = k$, then we call this maximum a *maximal k-compressible* set. A maximal 3-compressible set of B_5 can be obtained from our 3-compressible set B of the previous paragraph by adding three signed permutations: $(1, -5, 2, 3, 4)$, $(1, -4, -3, -2, 5)$, and $(1, 5, -4, -3, -2)$.

A maximal k-compressible set in B_n has an interesting culinary interpretation as the collection of all stacks of n burned pancakes having the property that k previously chosen pairs of adjacently sized pancakes are stuck together. Specifically, for each of the k pairs, the smaller pancake's burned side is stuck to the larger pancake's non-burned side, causing the pair to behave like a single pancake. Under this view, the extended pancakes at 0 and n are both one-sided: the former has only a burned side, while the latter has only a non-burned side.

Our third fact should hopefully come as no surprise, and it will be very useful in the next section.

Fact 3. *Let B be a maximal k-compressible set of B_n. Then the subgraph of SR_n (SP_n) induced by B is isomorphic to SR_{n-k} (SP_{n-k}).*

It would be nice if we had an analogue of k-compressibility leading to an isomorphism between unsigned graphs, but unfortunately we do not. For example, consider the set $S = \{(1, 2, 3), (2, 1, 3)\}$. We would like to dub S a "maximal 2-compressible" subset of S_3 with $J = \{1, 3\}$, but there is no use in doing so: the subgraph of UP_3 induced by S is certainly not isomorphic to UP_1, as these graphs have different orders. The source of the problem here is that the elements 1 and 2, when glued together, inherit an orientation.

3. Girth of pancake graphs

Before continuing, let us note that we may be led to suspect that pancake graphs might be bipartite. However, in the signed case, a cycle of length 9 in SP_4 is given by

$$(\sigma_{[4]}, \sigma_{[2]}, \sigma_{[3]}, \sigma_{[4]}, \sigma_{[3]}, \sigma_{[4]}, \sigma_{[1]}, \sigma_{[2]}, \sigma_{[1]})$$

which embeds in both SP_n and SR_n for $n \geq 4$ by Facts 1 and 2. As it turns out, this is the shortest odd cycle length in SP_n.

As for the unsigned case, the authors in [8] showed how to embed cycles of every length k in UP_n satisfying $6 \leq k \leq n!$, except for $k = n! - 1$. Their work immediately gives cycles of those lengths in UR_n by Fact 1. Here we show that no shorter cycles are possible.

First, we provide without proof a vital and commonly appearing fact from group theory. This lemma will be useful in finding the girth of SP_n and SR_n, and it is in fact, all we need to find the girth of UP_n and UR_n.

Lemma 4. *If $(\rho_1, \rho_2, \ldots, \rho_k)$ is a cycle of reversals, then for all $j \in [k]$, the cyclic shift*

$$(\rho_j, \rho_{j+1}, \ldots, \rho_k, \rho_1, \rho_2, \ldots, \rho_{j-1})$$

is a cycle as well.

Theorem 5. *UP_n has girth 6 for $n \geq 3$.*

Proof. First note that since UP_3 is a 2-regular connected graph of order 6, $UP_3 = C_6$. Also, because UP_m embeds in UP_n for $m < n$ by Fact 2, we automatically have that for $n \geq 3$, UP_n has girth at most 6.

So, assume that for some $n > 3$, UP_n has a cycle $(\rho_1, \rho_2, \ldots, \rho_k)$ of length $k < 6$. All the ρ_i are in the form $\upsilon_{[m_i]}$, so set $m = \max m_i$; by Lemma 4, we are free to set $\rho_1 = \upsilon_{[m]}$. Now, observe that since $\rho_1(m) = 1$, the only way that $\rho_k \cdots \rho_1$ can fix 1 is if some other ρ_i equals $\upsilon_{[m]}$. By the definition of a cycle, $\rho_2 \neq \rho_{[m]}$, so appealing again to Lemma 4, we are free to set $\rho_3 = \upsilon_{[m]}$ since $k < 6$; furthermore, no other ρ_i may equal $\upsilon_{[m]}$. Note that the fact that $\rho_3 = \upsilon_{[m]}$ immediately implies that $k \neq 3$ by the definition of cycle.

Also, $\rho_3(m) = \upsilon_{[m]}(m) = 1$, and we can see that $(\rho_3\rho_2\rho_1)(m) = (\upsilon_{[m]}\rho_2\upsilon_{[m]})(m)$ must equal m. Putting these two facts together gives $(\rho_2\upsilon_{[m]})(1) = m$. But since $\upsilon_{[m]}(1) = m$, we must have that ρ_2 fixes 1. This is impossible since ρ_2 is a prefix reversal, and so we have the desired contradiction. $\square$

Theorem 6. *UR_n has girth 4 for $n \geq 3$.*

Proof. First we note that UR_3 has a cycle of length 4, given by $(\upsilon_{[2,3]}, \upsilon_{[3]}, \upsilon_{[2,3]}, \upsilon_{[2]})$. Thus we only need to show that UR_n is triangle-free for all $n \geq 3$.

Once more, we proceed by contradiction. Assume that for some $n \geq 3$, (ρ_1, ρ_2, ρ_3) is a cycle of length 3 in UR_n, and note that by the definition of a cycle, the ρ_i must be distinct. Assume that $\rho_i = \upsilon_{[j_i,k_i]}$ for each $i = 1, 2, 3$, and set $j = \min j_i$. By Lemma 4, we may assume without loss of generality that $\rho_1 = \upsilon_{[j,k_1]}$.

We now branch off into three cases. In the first case, assume that $j_2 = j_3 = j$, so that the k_i are distinct. Setting $k = \max k_i$, we can assume without loss of generality that $\rho_1 = \upsilon_{[j,k]}$, and observe that while ρ_2 and ρ_3 fix k, $\rho_1(k) = j$, a clear contradiction to the fact that (ρ_1, ρ_2, ρ_3) is a triangle.

Second, if the j_i are distinct, then by the minimality of j both ρ_2 and ρ_3 would both fix j. Since $\rho_3\rho_2\rho_1 = I_n$, this forces ρ_1 to also fix j, which is also impossible.

Finally, if exactly one of j_2 and j_3 is equal to j, then by Lemma 4, we can assume without loss of generality that $j_2 = j$ and $j_3 > j$. Since $j_1 = j_2$, by the definition of a cycle, we must have that $k_2 \neq k_1$. Furthermore, because $j_3 > j$, we must have that ρ_3 fixes j, and so $(\rho_2\rho_1)(j) = j$. Yet $\rho_2(j) = k_2$, so that if $k_2 > k_1$, then

$$(\rho_2\rho_1)(j) = \rho_1(k_2) = k_2 \neq j.$$

On the other hand, if $k_2 < k_1$, then

$$(\rho_2\rho_1)(j) = \rho_1(k_2) = j + k_1 - k_2 \neq j.$$

Hence, in all cases we have obtained a contradiction to our assumption that UR_n contains a triangle. □

Now we turn to the slightly trickier case of signed pancake graphs. The fuel for finding the girth of these graphs is found in the following lemma. It will guarantee that in a short cycle, we cannot accumulate very many breakpoints, since in order to return to the identity, these breakpoints must be eliminated. Observe that it applies to both signed and unsigned cases.

Lemma 7. *For any reversal ρ and permutation α, there are at most two locations in α where ρ may create or remove a breakpoint.*

Proof. We will prove the signed case only, but there is an unsigned case which proceeds almost identically. Assume that $\beta \in B_n$ and that $\rho = \sigma_{[j,k]}$. For $0 \leq i < j - 1$ or $k < i \leq n$, we can see that $(\rho\beta)(i) = \beta(i)$, so that for these values of i, $\rho\beta$ has a breakpoint at i iff β has a breakpoint at i.

Furthermore, if i is such that $j - 1 < i < k$, then observe that $(\rho\beta)(i) = \beta(k - i + j)$ and $(\rho\beta)(i + 1) = \beta(k - (i + 1) + j)$, so that

$$(\rho\beta)(i + 1) - (\rho\beta)(i) \neq 1 \iff \beta(k - i + j) - \beta(k - (i + 1) + j) \neq 1.$$

This implies that for these values of i, $\rho\beta$ will have a breakpoint at i iff β has a breakpoint at $k - (i + 1) + j$. Note that since $j - 1 < i < k$, we have $j - 1 < k - i + j < k$ as well.

This leaves only two remaining possibilities: $i = j - 1$ and $i = k$. □

Our first corollary to Lemma 7 follows immediately.

Corollary 8. *A prefix reversal can create or remove at most one non-initial breakpoint from a permutation.*

The second corollary to Lemma 7 will allow us to quickly find the girth of SP_n.

Corollary 9. *If $(\rho_1, \rho_2, \ldots, \rho_k)$ is a cycle of prefix reversals, then for any $j \in [k]$, $\rho_j\rho_{j-1}\cdots\rho_1$ has at most $\min\{j, \lfloor k/2 \rfloor\}$ non-initial breakpoints. Furthermore, if the ρ_i are signed prefix reversals, then $B = \bigcup_{j=1}^{k}\{\rho_j\rho_{j-1}\cdots\rho_1\}$ constitutes an $(n - \lfloor k/2 \rfloor)$-compressible set of B_n.*

Proof. The first statement follows immediately from Corollary 8 and the Pigeonhole Principle. As for the second statement, it is trivial that B is $(n - \lceil k/2 \rceil)$-compressible. To strengthen this, assume that for some odd k, B is not $(n - \lfloor k/2 \rfloor)$-compressible. By the first part of this corollary, this implies that both $\rho_{\lfloor k/2 \rfloor}\cdots\rho_1$ and $\rho_{\lceil k/2 \rceil}\cdots\rho_1$ have $\lfloor k/2 \rfloor$ non-initial breakpoints. Since B is not $(n - \lfloor k/2 \rfloor)$-compressible, the two sets of breakpoints cannot be equal. But this means that $\rho_{\lceil k/2 \rceil}$ both creates and removes a non-initial breakpoint from $\rho_{\lceil k/2 \rceil}\cdots\rho_1$, which is impossible by Corollary 8. □

Theorem 10. *SP_n has girth 8 for $n \geq 2$.*

Proof. First, SP_2 is a 2-regular connected graph of order 8, so $SP_2 = C_8$; furthermore, SP_m embeds in SP_n for $m < n$, so for $n \geq 2$, SP_n contains a cycle of length 8.

Assume that for some $n > 2$, SP_n has a cycle $(\rho_1, \rho_2, \ldots, \rho_k)$ of length $k < 8$. By Corollary 9, the $\rho_j\rho_{j-1}\cdots\rho_1$ are $(n-3)$-compressible, so by Fact 3, we may assume without loss of generality that $n = 3$. A computer check verifies that SP_3 has no cycle of length less than 8. □

Theorem 11. *SR_n has girth 4 for $n \geq 2$.*

Proof. We first note that SR_2 is isomorphic to the discrete 3-dimensional hypercube and is therefore bipartite. Furthermore, SR_2 has girth 4; for an example of a 4-cycle in SR_2, consider $(\sigma_{[2,2]}, \sigma_{[1]}, \sigma_{[2,2]}, \sigma_{[1]})$.

So, assume that for some $n \geq 2$, SR_n has a triangle (ρ_1, ρ_2, ρ_3), and let us set $B = \{\rho_1, \rho_2\rho_1, \rho_3\rho_2\rho_1\}$. Now, ρ_1 must have two breakpoints. Furthermore, ρ_3 must remove two breakpoints from $\rho_2\rho_1$ to give the identity. Therefore, $\rho_2\rho_1$ has two breakpoints, and ρ_2 must create and remove one breakpoint from ρ_1. This implies that there are only three indices i, $0 \leq i \leq n$, such that any of the three members β of B has a breakpoint between the occurrence of i and $i + 1$ in β; letting J equal the set of these three indices, we can conclude that the set $\{\rho_1, \rho_2\rho_1, \rho_3\rho_2\rho_1\}$ is $(n - 2)$-compressible. Hence, by Fact 3, we may assume without loss of generality that $n = 2$. But we have already noted that the girth of SR_2 is 4. □

4. Conclusion

It is certainly interesting that pancake graphs should have constant girth. After all, the Moore bound for each family of pancake graphs is a quadratic or cubic polynomial in n, whereas the degrees of pancake graphs grow much faster, at $n!$ or $2^n \cdot n!$.

Considering how breakpoints have been used, without reference to other results, to find the girth of pancake graphs in this note, we wonder if this method could be strengthened to determine the girth of a larger class of Cayley graphs on permutation groups.

Finally, in the introduction we mentioned the potential biological value of pancake graphs. In this light, it is somewhat disappointing that pancake graphs have such small girth, since phylogeneticists religiously use parsimony to assume that evolution "takes the shortest path" when mutating one species into another. The presence of short cycles in pancake graphs slightly weakens this premise.

Acknowledgments

The author would like to thank Dr. Laurie Heyer for her guidance, the referees for insightful comments, and the Davidson Research Initiative for its support.

References

[1] V. Bafna, P.A. Pevzer, Sorting by transpositions, SIAM Journal on Discrete Mathematics 11 (1998) 224–240.
[2] A. Caprara, Sorting by reversals is difficult, in: Proceedings of the First Annual International Conference on Computational Molecular Biology, 1997, pp. 75–83.
[3] B. Chitturi, W. Fahle, Z. Meng, L. Morales, C.O. Shields, I.H. Sudborough, W. Voit, An $(18/11)n$ upper bound for sorting by prefix reversals, Theoretical Computer Science 410 (2009) 3372–3390.
[4] D.S. Cohen, M. Blum, On the problem of sorting burnt pancakes, Discrete Applied Mathematics 61 (1995) 105–120.
[5] G. Fertin, A. Labarre, I. Rusu, E. Tannier, S. Vialette, Combinatorics of Genome Rearrangements, MIT Press, 2009.
[6] W. Gates, C. Papadimitriou, Bounds for sorting by prefix reversals, Discrete Mathematics 27 (1979) 47–57.
[7] S. Hannenhalli, P.A. Pevzner, Transforming cabbage into turnip: polynomial algorithm for sorting signed permutations by reversals, Journal of the Association for Computing Machinery 46 (1999) 1–27.
[8] A. Kanevsky, C. Feng, On the embedding of cycles in pancake graphs, Parallel Computing 21 (1995) 923–936.
[9] J. Kececioglu, D. Sankoff, Exact and approximation algorithms for sorting by reversals, with application to genome rearrangement, Algorithmica 13 (1995) 180–210.
[10] P.A. Pevzner, Computational Molecular Biology: An Algorithmic Approach, first ed., MIT, Cambridge, Massachusetts, 2000.