Intersection reverse sequences and geometric applications

Adam Marcus* Gábor Tardos†

Abstract

Pinchasi and Radoičić [11] used the following observation to bound the number of edges of a topological graph without a self-crossing cycle of length 4: if we make a list of the neighbors for every vertex in such a graph and order these lists cyclically according to the order of the emanating edges, then the common elements in any two lists have reversed cyclic order. Building on their work we give an improved estimate on the size of the lists having this property. As a consequence we get that a topological graph on n vertices not containing a self-crossing C_4 has $O(n^{3/2}\log n)$ edges. Our result also implies that n pseudo-circles in the plane can be cut into $O(n^{3/2}\log n)$ pseudo-segments, which in turn implies bounds on point-curve incidences and on the complexity of a level of an arrangement of curves.

1 Introduction

In this paper we consider *cyclically ordered sequences* of distinct symbols from a finite alphabet. We say that two such sequences are *intersection reverse* if the common elements appear in reversed cyclic order in the two sequences. A collection of cyclically ordered sequences $s_1, s_2, \ldots, s_m$ will be referred to as *pairwise intersection reverse* if the sequences s_i and s_j are intersection reverse for all $1 \leq i < j \leq m$.

A *topological graph* is a graph without loops or multiple edges drawn in the plane (vertices correspond to distinct points, edges correspond to Jordan curves connecting the corresponding vertices). We assume no edge passes through a vertex other than its endpoints and every two edges have a finite number of common interior points and they properly cross at each of these points. For a vertex v of a topological graph G let $L_G(v)$ be the list of its neighbors ordered

*Visiting Researcher, Alfréd Rényi Institute, 1364 Budapest, Pf.127, Hungary, on leave from the Department of Mathematics (ACO), Georgia Institute of Technology, Atlanta, GA 30332-0160. Email: adam@math.gatech.edu. This research was made possible due to funding by the Fulbright Program in Hungary.

†Alfréd Rényi Institute, 1364 Budapest, Pf.127, Hungary. Email: tardos@renyi.hu. Partially supported by Hungarian National Research Fund grants OTKA T048826, T046234, T037846.

cyclically counterclockwise according to the initial segment of the connecting edge.

Pinchasi and Radoičić [11] noticed the following simple fact:

Fact 1. *If the lists $L_G(u)$ and $L_G(v)$ are not intersection reverse for two distinct vertices u and v of the topological graph G, then G contains a self-crossing cycle of length 4. Moreover, u and v are opposite vertices of a cycle of length 4 in G that has two edges crossing an odd number of times.*

For the proof one only has to consider drawings of the complete bipartite graph $K_{2,3}$ (see details in [11]). Pinchasi and Radoičić used Fact 1 to bound the number of edges of a topological graph not containing a self-crossing C_4. They showed that such a graph on n vertices has $O(n^{8/5})$ edges. Following in their footsteps, we use the same property to improve their bound to $O(n^{3/2}\log n)$. This bound is tight apart from the logarithmic factor since there exist (abstract) simple graphs on n vertices with $\Omega(n^{3/2})$ edges containing no C_4-subgraph (see, for example, [7]). Our main technical result is the following:

Theorem 1. *Let $A^1, A^2, \ldots, A^m$ be a collection of cyclically ordered lists, each containing a d-element subset of a set of n symbols. If these lists are pairwise intersection reverse, then*

$$d = O\left(\sqrt{n}\log n + \frac{n}{\sqrt{m}}\right).$$

We give the proof of this theorem in Section 2. In Section 3 we present its consequences, among them the bound on the number of edges in any topological graph that does not contain a self-crossing C_4.

The most important consequence of Theorem 1 deals with collections of *pseudo-circles*: simple closed Jordan curves, any two of which intersect at most twice, with proper crossings at each intersection. The result readily generalizes to unbounded open curves such as *pseudo-parabolas*, the graphs of continuous real functions defined on the entire real line such that any two intersect at most twice and they properly cross at these intersections.

Tamaki and Tokuyama [12] were the first to consider the problem of *cutting* pseudo-parabolas into *pseudo-segments*, i.e., subdividing the original curves into segments such that any two segments intersect at most once. Such a separation turns out to be quite useful since pseudo-segments are much easier to work with than pseudo-parabolas and pseudo-circles, as will be seen in Section 3.

Tamaki and Tokuyama [12] proved that n pseudo-parabolas can be cut into $O(n^{5/3})$ pseudo-segments. This was extended to x-monotone pseudo-circles by Aronov and Sharir [3] and by Agarwal, et al. [2]. It was also improved for certain collections of curves that admit a three-parameter algebraic parameterization to $n^{3/2}\log^{\alpha^{O(1)}(n)}(n)$, where α is the inverse Ackermann function.

Previously, the best bound on the number of cuts needed for arbitrary collections of pseudo-parabolas and x-monotone pseudo-circles was $O(n^{8/5})$ [2], which uses the result of Pinchasi and Radoičić on topological graphs without a self-crossing C_4. With our improvement of the latter result, we can prove that n pseudo-parabolas can be cut into $O(n^{3/2} \log n)$ pseudo-segments. This substantially improves the previous bounds for arbitrary collections and is still slightly better than results on families with algebraic parameterization; we reduce a factor which grows slightly faster than polylogarithmically to a single logarithmic factor. In doing so, we are able to simplify the results in [2, 11, 12], as well as generalize them to the cases when the pseudo-parabolas and pseudo-circles are not necessarily x-monotone.

In Section 3 we show the above result, as well as its applications to point-curve incidences and the level complexities of curve arrangements. See [1, 2, 3, 4, 5, 12] for more details and applications.

Finally in Section 4 we discuss a few related problems that are still open.

All logarithms in this paper are binary.

2 Intersection reverse sequences

In this section we prove our main technical result, Theorem 1. Much of the proof follows the argument of Pinchasi and Radoičić [11]. We start with an overview of their techniques and comment on similarities and differences with the present proof.

Pinchasi and Radoičić break the cyclically ordered lists into linearly ordered blocks. They consider pairs of blocks from separate lists and pairs of symbols contained in both blocks. They distinguish between *same pairs* and *different pairs* according to whether the two symbols appear in the same or in different order. They observe that any pair of symbols that appears in many blocks must produce almost as many same pairs as different pairs. On the other hand the intersection reverse property forces two cyclically ordered lists—unless most of their intersection is concentrated into a single pair of blocks—to contribute many more different than same pairs. Exceptional pairs of cyclically ordered lists are treated separately with techniques from extremal graph theory. They optimize in their choice for the length of the blocks.

We follow almost the same path, but instead of optimizing for block length we consider many block lengths (an exponential sequence) simultaneously. For two intersection reverse lists, no block length yields significantly more same pairs than different pairs. On the other hand, we will show that at least one of the block lengths actually gives many more different pairs than same pairs. As a consequence we do not have to bound "exceptional pairs" of lists separately.

Definition. We will use the term *sequence* to denote a linearly ordered list of distinct symbols and the term *cyclic sequence* to denote a cyclically ordered list of distinct symbols. Clearly, if one breaks up a cyclic sequence into blocks, then the blocks are (linearly ordered) sequences. For a sequence or cyclic sequence A we write $\overline{A}$ for the set of symbols in A. We define intersection reverse for sequences just as for cyclic sequences: we say that the sequences A and B are *intersection reverse* if they induce inverse linear orders on $\overline{A} \cap \overline{B}$. If two sequences are not intersection reverse, we call them *singular*. Note that if two sequences A and B have $|\overline{A} \cap \overline{B}| \leq 1$, then the sequences are trivially intersection reverse. The same holds for cyclic sequences A and B if $|\overline{A} \cap \overline{B}| \leq 2$.

For a sequence B and symbols $a \neq b$ we define

$$f(B,a,b) = \begin{cases} 0 & \text{if } a \notin \overline{B} \text{ or } b \notin \overline{B}, \\ 1 & \text{if } a \text{ precedes } b \text{ in } B, \\ -1 & \text{if } b \text{ precedes } a \text{ in } B. \end{cases}$$

For two sequences B and B' we let

$$f(B,B',a,b) = f(B,a,b)f(B',a,b).$$

Notice that $f(B,B',a,b) = 1$ for same pairs and $f(B,B',a,b) = -1$ for different pairs, and that $\sum f(B,B',a,b)$ corresponds to the difference between the number of same pairs and different pairs.

The next lemma is taken from [11]. We will use the notation $\sum_{a \neq b}$ (both here and later in this section) to denote a sum taken over all ordered pairs of distinct symbols a and b.

Lemma 2. *Let the cyclic sequences A and A' consist of the (linearly ordered) blocks $B_1, \dots, B_k$ and $B'_1, \dots, B'_{k'}$, respectively. If A and A' are intersection reverse, then at most one of the pairs B_i, B'_j is singular. For this singular pair we have*

$$\sum_{a \neq b} f(B_i, B'_j, a, b) \leq |\overline{B_i} \cap \overline{B'_j}|.$$

For all of the other (intersection reverse) pairs B_i, B'_j we have

$$\sum_{a \neq b} f(B_i, B'_j, a, b) = |\overline{B_i} \cap \overline{B'_j}| - |\overline{B_i} \cap \overline{B'_j}|^2.$$

Proof. Let B_i and B'_j be blocks with common symbols appearing in the order $a_1, \dots, a_l$ in B_i. Due to the intersection reverse property of A and A', they appear in the order $a_x, a_{x-1}, \dots, a_1, a_l, a_{l-1}, \dots, a_{x+1}$ in B'_j for some $1 \leq x \leq l$. Note that B_i and B'_j are singular if and only if $x < l$, and it is easy to verify that this can happen for at most a single pair of blocks. For a singular pair, we have

$$\begin{aligned} \sum_{a \neq b} f(B_i, B'_j, a, b) &= [2x(l-x)] - [x(x-1) + (l-x)(l-x-1)] \\ &= l - (l-2x)^2 \leq l. \end{aligned}$$

For all intersection reverse pairs, however, all pairs of symbols $a \neq b$ from the intersection $\overline{B_i} \cap \overline{B'_j}$ contribute -1 to the sum. □

For the rest of the section, assume that we have the collection of pairwise intersection reverse cyclic sequences $A^1, \dots, A^m$ from the theorem (recall that each consists of a d-element subset of a set of n symbols). Also, let $p^{ij} = |\overline{A^i} \cap \overline{A^j}|$, and $p = \sum_{i \neq j} p^{ij}$. First we bound p based on the limited size of the alphabet. For simplicity we assume $dm > 2n$ (otherwise Theorem 1 is immediate).

Lemma 3. $p \geq \frac{d^2m^2}{2n}$ *and* $\sum_{i \neq j}(p^{ij})^2 \geq \frac{p^2}{m^2}$.

Proof. Let d_a be the number of times the symbol a appears among the cyclic sequences A^i. Then a contributes $d_a^2 - d_a$ to p, so we have $p = \sum_a d_a^2 - \sum_a d_a$, where the summation is over the n different symbols a. We also have $\sum_a d_a = dm$ as it is the sum of the sizes of the sequences A^i. Applying the inequality between the quadratic and the arithmetic mean and using $dm > 2n$ we obtain

$$p = \sum_a d_a^2 - \sum_a d_a \geq \frac{1}{n}\left(\sum_a d_a\right)^2 - \sum_a d_a = \frac{d^2m^2}{n} - dm \geq \frac{d^2m^2}{2n}.$$

The second inequality in the lemma is also due to the inequality between the quadratic and arithmetic means, as

$$\sum_{i \neq j}(p^{ij})^2 \geq \frac{1}{m^2 - m}\left(\sum_{i \neq j}(p^{ij})\right)^2 > \frac{p^2}{m^2}.$$

□

We now split each A^i into two almost equal size consecutive blocks A_0^i and A_1^i. In general, for a 0–1 sequence s we split the block A_s^i of A^i into two almost equal halves (differing in size by at most 1): A_{s0}^i and A_{s1}^i. The cyclic order of A^i linearly orders the elements in each of these blocks. Let $k = \lceil \log d \rceil < \log n + 1$. Clearly, any 0–1 sequence s of length k satisfies $|\overline{A_s^i}| \leq 1$.

For $1 \leq i \leq m$ and $1 \leq j \leq m$ we let

$$S^{ij} = \sum_{l=1}^{k} w_l \sum_{\substack{a \neq b \\ |s|=|t|=l}} f(A_s^i, A_t^j, a, b),$$

where the outer summation is taken over lengths $1 \leq l \leq k$ and the inner summation is taken over all pairs of symbols $a \neq b$ and all 0–1 sequences s and t of size $|s| = |t| = l$. We consider the pair (a, b) to be ordered, thereby double counting each unordered pair. The weights w_l in the formula are positive and we set them later. Our goal is to contrast a lower bound on $\sum_{i \neq j} S^{ij}$ (or rather on the partial sum for fixed symbols $a \neq b$) with upper bounds on the individual S^{ij}. Again we consider the (i, j) pairs to be ordered, resulting in another double counting.

The lower bound is straightforward:

Lemma 4. $\sum_{i \neq j} S^{ij} \geq -md^2 \sum_{l=1}^{k} \frac{w_l}{2^l}$.

Proof. Notice that for fixed a, b, and l we get a perfect square when summing over all i and j. In particular,

$$\sum_{i=1}^{m}\sum_{j=1}^{m} S^{ij} = \sum_{l=1}^{k} w_l \sum_{a\neq b}\left(\sum_{i=1}^{m}\sum_{|s|=l} f(A_s^i, a, b)\right)^2 \geq 0$$

We can bound the S^{ii} terms separately as they are merely a (weighted) counting of the number of pairs contained in each block. Since $|\overline{A_s^i}| < d/2^{|s|}+1$, we have

$$\sum_{i\neq j} S^{ij} = \sum_{i=1}^{m}\sum_{j=1}^{m} S^{ij} - \sum_{i=1}^{m} S^{ii} \geq 0 - \sum_{i=1}^{m}\sum_{l=1}^{k} 2w_l \sum_{|s|=l}\binom{|\overline{A_s^i}|}{2}$$
$$\geq -md^2\sum_{l=1}^{k}\frac{w_l}{2^l}.$$

□

The upper bound, however, requires more effort.

Lemma 5. *For $i \neq j$ we have*

$$S^{ij} \leq p^{ij}\sum_{l=1}^{k} w_l - \frac{(p^{ij})^2}{4\sum_{l=1}^{k}\frac{1}{w_l}}.$$

Proof. We fix the indices $i \neq j$ and consider the following quantities:

- $r_{st} = |\overline{A_s^i} \cap \overline{A_t^j}|$ and
- $Q_{st} = \sum_{a\neq b} f(A_s^i, A_t^j, a, b)$

where s and t are 0–1 sequences of equal length.

For a fixed length $1 \leq l \leq k$, the blocks A_s^i with $|s| = l$ form a subdivision of A^i, while the blocks A_t^j with $|t| = l$ form a subdivision of A^j. By Lemma 2, there is at most one singular pair (A_s^i, A_t^j) for any fixed length $|s| = |t| = l$. For these singular pairs we have

$$Q_{st} \leq r_{st},$$

while for the intersection reverse ones we have

$$Q_{st} = r_{st} - r_{st}^2.$$

Recall that any pair of sequences of length at most 1 is intersection reverse, so we do not find any singular pairs when $|s| = |t| = k$.

For a 0–1 sequence s of length $|s| > 1$ let s' denote the sequence obtained from s by deleting its last digit, hence the block $A_{s'}^i$ contains the smaller block A_s^i. We call a pair (s, t) of equal length 0–1 sequences a *leader pair* if (A_s^i, A_t^j) is intersection reverse and either $|s| = |t| = 1$ or the pair $(A_{s'}^i, A_{t'}^j)$ is singular.

Since $(A^i_{s'}, A^j_{t'})$ is singular for at most one pair (s', t') of a fixed length, it follows that there can be at most 4 leader pairs (s,t) at the next bigger length. Furthermore, any symbol $a \in \overline{A^i} \cap \overline{A^j}$ appears in $\overline{A^i_s} \cap \overline{A^j_t}$ for exactly one leader pair (s,t): the longest intersection reverse pair of blocks containing them (recall that we only consider pairs of blocks with equal length subscripts). Thus we have $\sum_{(s,t)\in L} r_{st} = p^{ij}$ for the set L of leader pairs.

We use $Q_{st} = r_{st} - r_{st}^2$ for leader pairs (s,t) only. For all other pairs, intersection reverse or singular, we use $Q_{st} \le r_{st}$:

$$\begin{aligned} S^{ij} =& \sum_{l=1}^{k} w_l \sum_{|s|=|t|=l} Q_{st} \\ \le& \sum_{l=1}^{k} w_l \sum_{|s|=|t|=l} r_{st} - \sum_{(s,t)\in L} w_{|s|} r_{st}^2 \\ =& p^{ij} \sum_{l=1}^{k} w_l - \sum_{(s,t)\in L} w_{|s|} r_{st}^2 \end{aligned}$$

since $\sum_{|s|=|t|=l} r_{st} = p^{ij}$ for any fixed l. The Cauchy-Schwarz inequality gives

$$\left(\sum_{(s,t)\in L} w_{|s|} r_{st}^2 \right) \left(\sum_{(s,t)\in L} \frac{1}{w_{|s|}} \right) \ge \left(\sum_{(s,t)\in L} r_{st} \right)^2 = (p^{ij})^2.$$

Here $\sum_{(s,t)\in L} (1/w_{|s|}) \le 4 \sum_{l=1}^{k} (1/w_l)$, so

$$\sum_{(s,t)\in L} w_{|s|} r_{st}^2 \ge \frac{(p^{ij})^2}{4 \sum_{l=1}^{k} \frac{1}{w_l}}$$

and we conclude that

$$S^{ij} \le p^{ij} \sum_{l=1}^{k} w_l - \frac{(p^{ij})^2}{4 \sum_{l=1}^{k} \frac{1}{w_l}}$$

as claimed. □

Comparing the two estimates in the previous lemmas gives the theorem.

Proof of Theorem 1. Using Lemmas 4, 5, and 3 (respectively), we obtain

$$
\begin{aligned}
-md^2 \sum_{l=1}^{k} \frac{w_l}{2^l} \leq & \sum_{i \neq j} S^{ij} \\
\leq & \sum_{i \neq j} p^{ij} \sum_{l=1}^{k} w_l - \frac{\sum_{i \neq j} (p^{ij})^2}{4 \sum_{l=1}^{k} \frac{1}{w_l}} \\
\leq & p \sum_{l=1}^{k} w_l - \frac{p^2}{4m^2 \sum_{l=1}^{k} \frac{1}{w_l}}.
\end{aligned}
$$

This inequality implies that either $p \leq 8m^2(\sum_{l=1}^{k} w_l)(\sum_{l=1}^{k}(1/w_l))$ or $p^2 \leq 8d^2m^3(\sum_{l=1}^{k}(w_l/2^l))(\sum_{l=1}^{k}(1/w_l))$. By Lemma 3, we have that $p \geq d^2m^2/(2n)$, so either

$$
d \leq 4\sqrt{n} \sqrt{\sum_{l=1}^{k} w_l} \sqrt{\sum_{l=1}^{k} \frac{1}{w_l}}
$$

or

$$
d \leq \frac{6n}{\sqrt{m}} \sqrt{\sum_{l=1}^{k} \frac{w_l}{2^l}} \sqrt{\sum_{l=1}^{k} \frac{1}{w_l}}.
$$

We choose the weights w_l now. Equal weights ($w_l = 1$) yield $d = O(\sqrt{n} \log n + n\sqrt{\log n}/\sqrt{m})$, but we can improve on this bound by choosing

$$
w_l = \frac{1}{1 + \frac{k}{2^{l/2}}}.
$$

In this case $\sum_{l=1}^{k} w_l \leq k$, $\sum_{l=1}^{k}(1/w_l) \leq 4k$, and $\sum_{l=1}^{k}(w_l/2^l) \leq 3/k$. Thus we either have $d \leq 8k\sqrt{n}$ or $d \leq 21n/\sqrt{m}$ and the statement of the theorem follows. □

3 Consequences

In this section we present several geometric applications of Theorem 1.

3.1 Self-crossing cycles of length 4

Any bound for the $n = m$ case of Theorem 1 carries over to the number of edges of a topological graph not containing a self-crossing C_4 by [11]. Using the following corollary, however, the proof is even simpler:

Corollary 6. *Let us be given m cyclic sequences over an n-element set of symbols. If the cyclic sequences are pairwise intersection reverse, then the sum of their sizes is $O(m\sqrt{n} \log n + n\sqrt{m})$.*

Proof. Let c be the hidden constant in the statement of Theorem 1 (our proof gives $c = 21$) and define $t_k = c\sqrt{n}\log n + 2^k cn/\sqrt{m}$ for positive integers k (setting $t_0 = 0$). We define m_k to be the number of cyclic sequences whose lengths lie in the interval $(t_k, t_{k+1}]$. For $k \geq 1$, if we prune each of the m_k sequences to be exactly length t_k and apply the uniform result derived in the previous section, we get that $m_k \leq m/4^k$ (note this is trivially true for $k = 0$ as well). Thus we have that the sum of the lengths of the sequences is at most

$$\sum_{k=0}^{\infty} m_k t_{k+1} = cm\sqrt{n}\log n + c\left(\sum_{k=0}^{\infty} 2^{k+1} m_k\right)\frac{n}{\sqrt{m}} = O(m\sqrt{n}\log n + n\sqrt{m})$$

□

Corollary 7. *If an n-vertex topological graph does not contain a self-crossing C_4 it has $O(n^{3/2}\log n)$ edges. The same holds if every pair of edges in every C_4 subgraph cross an even number of times.*

Proof. The statements are direct consequences of Corollary 6 using Fact 1, since the sum of the sizes of the lists of neighbors is the sum of the degrees, i.e., twice the number of edges. □

3.2 Cutting Number

Tamaki and Tokuyama [12] considered the *cutting number* of a collection of curves. This is defined to be the least number of cuts needed to obtain a collection of shorter curves, each pair of which intersects at most once. This was in turn shown to be directly related to Corollary 7 in [2].

The restriction of the next corollary to so called *x-monotone* pseudo-circles can be derived from Corollary 7 using the combination of techniques in the papers [2, 12]. Here we give a simple and direct argument that does not require any additional monotonicity assumption on the pseudo-circles. Recall that this result slightly improves the best previous bound for (x-monotone) pseudo-circles with a three parameter algebraic representation as defined in [2] (such as ordinary circles) and substantially improves the previous bounds for pseudo-circles lacking such representation. For the definition of pseudo-circles see Section 1.

Corollary 8. *An arrangement of n pseudo-circles can be cut at $O(n^{3/2}\log n)$ points such that the resulting curves form a system of pseudo-segments.*

Before proving this result we define a few useful concepts related to pseudo-circles.

Definition. A simple closed Jordan curve (such as a pseudo-circle) cuts the plane into two open regions. We call the bounded region the *interior* of the pseudo-circle. Following [12] we define a *lens* to be the union of two segments from distinct pseudo-circles if they form a closed curve. The two segments constituting the lens are called the *sides* of the lens. A side of a lens is *positive* if the interior of the corresponding pseudo-circle contains the other side of the lens. A lens is classified as a *lens-face* if both sides are positive, a *moon-face* if it has a positive and a negative side, and an *inverse-face* if it contains two negative sides. We will also consider each pseudo-circle itself to be a (degenerate) lens. A collection of *non-overlapping lenses* is a set of lenses such that no segment of any pseudo-circle is contained in more than one lens. The different types of non-degenerate lenses are illustrated in Figure 1.

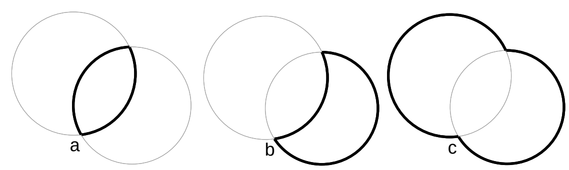

Figure 1: Examples of (a) a lens-face, (b) a moon-face, and (c) an inverse-face.

Notice that non-overlapping lenses may cross each other. For a collection $\mathcal{C}$ of pseudo-circles we let $\nu(\mathcal{C})$ denote the maximum size of a non-overlapping family of lenses and $\tau(\mathcal{C})$ denote the minimum number of cuts that transforms $\mathcal{C}$ into a collection of pseudo-segments. We do not allow for cuts at intersection points of the curves. The following lemma first appeared in [12], however, our proof takes a different approach. Apart from being shorter, it can also be easily extended to collections of curves which are allowed to intersect more than twice.

Lemma 9. $\tau(\mathcal{C}) = O(\nu(\mathcal{C}))$.

Proof. We consider the lenses as a hypergraph: the vertices are the segments of the pseudo-circles connecting adjacent intersection points, the edges are the collections of these segments forming a lens. With this notation $\nu(\mathcal{C})$ is the *packing (or matching) number* of this hypergraph, i.e., the maximum number of pairwise disjoint edges. Similarly, $\tau(\mathcal{C})$ is the *transversal (or piercing) number* of the hypergraph, i.e., the minimum size of a collection of vertices that intersects every edge. After the cuts, the resulting curves will form a system of pseudo-segments if and only if we cut every lens at least once. We always have $\tau(H) \geq \nu(H)$ for any hypergraph H, and much research has been focused on the connection between the packing and the transversal numbers. Tamaki and Tokuyama use a general result of Lovász [10] connecting these numbers to deduce their bound. We use, instead, the more specific result $\tau = O(\nu)$ for the families of so called 2-intervals (a 2-interval is simply a union of two intervals of the real line). This was proved by Tardos [13], and later Kaiser [8] proved the tight bound $\tau \leq 3\nu$. Our lenses are almost 2-intervals: they consist of two intervals, but of pseudo-circles (not the real line). We start by cutting every pseudo-circle at an arbitrary point. Now our pseudo-circles can be identified with disjoint intervals of the real line. With this identification, all lenses which remain after the first set of cuts correspond to 2-intervals (a disjoint union of the two sides of the lens). Using Kaiser's result we have $\tau(\mathcal{C}) \leq 3\nu(\mathcal{C}) + n$, where n is the number of pseudo-circles. Clearly $n \leq \nu(\mathcal{C})$ as the collection of degenerate lenses is non-overlapping, so we have $\tau(\mathcal{C}) \leq 4\nu(\mathcal{C})$ and this finishes the proof. □

Lemma 9 and the following lemma prove Corollary 8.

Lemma 10. *A collection of non-overlapping lenses in an arrangement of n pseudo-circles has $O(n^{3/2}\log n)$ lenses.*

Proof. Given an arrangement $\mathcal{C}$, let L be a set of non-overlapping lenses with L^{lens}, L^{moon}, and L^{inv} the sets of lens-faces, moon-faces, and inverse-faces in L (respectively). It is enough to prove the bound separately for each of these subsets, since the total number of degenerate lenses is only n.

For each $c \in C$, and each subset L^k (for $k = lens, moon, inv$) we make a list S_c^k consisting of all pseudo-circles $c' \in \mathcal{C}$ that form a lens in L^k together with c. For the lenses in L^{moon}, however, we include c' in the list S_c^{moon} only if the corresponding lens has its positive side in c and its negative side in c' (otherwise it will appear in $S_{c'}^{moon}$).

We then order each of the lists S_c^k according to the counterclockwise cyclic order of these lenses around c. Since all of the lenses are non-overlapping, this cyclic order is well defined.

The main observation is that, for fixed $k \in \{inv, lens, moon\}$, the lists S_c^k must be pairwise intersection reverse. As in the proof of Fact 1, one can prove this observation by considering the arrangements of 5 pseudo-circles forming six non-overlapping lenses. Notice that there are only a finite number of combinatorially different arrangements of 5 pseudo-circles in the plane. Instead of the simple but tedious case analysis we present three "counterexamples" where three pseudo-circles appear in the same cyclic order in the lists S_a and S_b. Here a and b are two of the pseudo-circles and we let S_a (respectively S_b) be the cyclic list of all pseudo-circles that together with a (respectively with b) form a lens in L (see Figure 2). Considering the lists S_a^{lens}, S_a^{moon} and S_a^{inv} separately resolves the problem. In the first example, for the lists S_a we had to consider two lens-faces and a moon-face from L, while in the second example for S_a we considered moon-faces and for S_b we considered lens-faces. For the third example we considered only moon-faces in L, but the moon-faces considered for S_a have their negative (rather than positive) side on a.

By Corollary 6, the sum of the length of the lists S_c^k is $O(n^{3/2}\log n)$ for each k. Hence the sum of the lengths of all of the lists is $O(n^{3/2}\log n)$ as well—but this sum is at least the size of L. □

Corollary 8 naturally generalizes to collections of open Jordan curves including, for example, pseudo-parabolas. We call a collection of simple closed and open Jordan curves a *generalized pseudo-circle collection* if both ends of every open curve are at infinity, any two curves have at most two points of intersection, and the curves cross properly at each intersection.

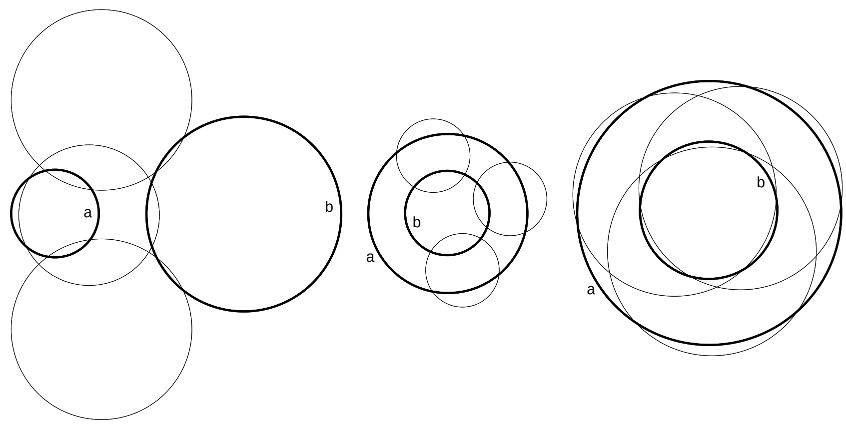

Figure 2: Three "counterexamples" to the intersection reverse property of S_a and S_b.

Corollary 11. *A generalized pseudo-circle collection $\mathcal{C}$ of n curves can be cut at $O(n^{3/2} \log n)$ points such that the resulting curve segments form a system of pseudo-segments.*

Proof. Given $\mathcal{C}$, we turn the arrangement into a system of n pseudo-circles and apply Corollary 8. Since there are a finite number of intersections, there is a sufficiently large circle D which contains all of them, together with all closed curves and all the segments of the open curves connecting two intersection points.

We modify the open curves in $\mathcal{C}$ outside the circle D by closing them. We can choose the arcs closing up the open curves in such a way that any two of the curves intersect at most once outside D. Therefore any pair in the resulting family $\mathcal{C}'$ intersects at most 3 times in total. Furthermore, $\mathcal{C}'$ consists of closed curves with proper intersections, so any pair of them must cross an even number of times. Thus $\mathcal{C}'$ is, in fact, a collection of pseudo-circles and Corollary 8 finishes the proof. □

3.3 Levels

Corollary 8 also has many consequences in the study of *levels* in arrangements of curves. Tamaki and Tokuyama [12] were first to show the usefulness of cutting numbers in this area, and progress has been made by Chan [4, 5].

Definition. Let $\mathcal{C}$ be the set of points in the graphs of the real functions $f_1, f_2, \ldots, f_n$. We assume that each f_i is continuous and defined everywhere on the real line, and that any pair of curves in $\mathcal{C}$ intersects a finite number of times. We define the k^{th} *level* of $\mathcal{C}$ to be the closure of the locus of points (x, y) on the curves in $\mathcal{C}$ with $|\{i : f_i(x) \leq y\}| = k$. The k^{th} level consists of portions of the curves in $\mathcal{C}$, delimited by intersections between these curves. We will call the total number of curve segments in a level its *complexity*.

Chan [5] derives an upper bound on the complexity of a given level of a collection of pseudo-parabolas by recursively estimating the number of intersections that can appear within a range of levels. Our improved bound in Corollary 11 improves Chan's analysis. We sketch the reasoning below.

Let $\mathcal{C}$ be a collection of n pseudo-parabolas and fix k. Let t_i stand for the number of intersections strictly between levels $k-i$ and $k+i$. The main inequality (Lemma 3.1) in [5] asserts that

$$t_i \le 2i(t_{i+1} - t_i) + O(ni + \Lambda_i),$$

where Λ_i is the number of lenses (formed by the curves in $\mathcal{C}$) lying strictly between levels $k-i$ and $k+i$. Lemma 4.1 of the same paper bounds Λ_i:

$$\Lambda_i = O(i^2 \nu(n/i)),$$

where $\nu(k)$ stands for the number of cuts needed to turn k pseudo-parabolas into a collection of pseudo-segments. By our Corollary 11 we have $\nu(k) = O(k^{3/2} \log k)$.

Putting these three inequalities together gives the recurrence

$$t_i \le 2i(t_{i+1} - t_i) + O(i^{1/2} n^{3/2} \log n).$$

Using $t_n = O(n^2)$ and solving the recurrence yields a bound on t_2 and therefore on the complexity of the k^{th} level.

Corollary 12. *Let $\mathcal{C}$ be a collection of n pseudo-parabolas. Then the maximum complexity of any level of $\mathcal{C}$ is $O(n^{3/2} \log^2 n)$.*

The above corollary represents a substantial improvement over the previous bound of $O(n^{8/5})$ for an arbitrary collection of pseudo-parabolas in [5]. For a collection possessing a three-parameter algebraic representation (as defined in [2]) the improvement is marginal, replacing a term which grows slightly faster than polylogarithmically with the term $\log^2 n$. These improvements carry over to levels of arrangements of algebraic curves of degree higher than two by the technique of *bootstrapping*, as developed in [5]. We do not state these slightly improved bounds here.

3.4 Incidences and Faces

Aronov and Sharir [3] and Agarwal, Aronov and Sharir [1] used cutting numbers in their analysis of the relations between curves and points in the plane.

Definition. Let $\mathcal{C}$ be a set of curves and $\mathcal{P}$ a set of points in the plane. We define $I(\mathcal{C}, \mathcal{P})$ to be the number of *incidences* between $\mathcal{C}$ and $\mathcal{P}$, that is the number of pairs $(c, p) \in \mathcal{C} \times \mathcal{P}$ such that curve c contains point p. We also define $K(\mathcal{C}, \mathcal{P})$ to be the sum of the complexities of the faces in the arrangement $\mathcal{C}$ which contain at least one point in P (assuming now that they are not on the curves). Here a *face* is a connected component of the complement of the union of the curves in $\mathcal{C}$, and the *complexity* of a face is defined to be the number of curve segments that comprise its boundary.

The results in [2] relate the values of $I(\mathcal{C},\mathcal{P})$ and $K(\mathcal{C},\mathcal{P})$ to the cutting numbers $\tau(\mathcal{C})$ discussed above. The following bounds were shown:

Lemma 13. *If $\mathcal{C}$ is a collection of n curves and $\mathcal{P}$ is a set of m points, then*

$$I(\mathcal{C},\mathcal{P}) = O(m^{2/3}n^{2/3} + m + \tau(\mathcal{C})),$$

$$K(\mathcal{C},\mathcal{P}) = O(m^{2/3}n^{2/3} + m + \tau(\mathcal{C})\log^2 n).$$

Thus, by Corollary 11, we have

Corollary 14. *If $\mathcal{C}$ is a collection of n generalized pseudo-circles and $\mathcal{P}$ is a set of m points, then*

1. $I(\mathcal{C},\mathcal{P}) = O(m^{2/3}n^{2/3} + m + n^{3/2}\log n)$
2. $K(\mathcal{C},\mathcal{P}) = O(m^{2/3}n^{2/3} + m + n^{3/2}\log^3 n)$

For curves that admit a three-parameter algebraic representation (see [2]) Chan [5] is able to improve the incidence and complexity bounds in Corollary 14 by applying them separately to smaller subsets of the points and curves. Our results also improve these better bounds, but only marginally, and therefore we do not state them here.

4 Open problems

The results in this paper raise a number of interesting questions. Corollary 7 is tight except possibly for the logarithmic factor as graphs with n vertices and $\Omega(n^{3/2})$ edges are known which do not contain any C_4 (see, for example, [7]). This also implies that the special cases of Theorem 1 and Corollary 6 when $n = m$ are almost tight. Nevertheless, it would be interesting to know if the logarithmic factor is needed.

Problem 1. *Is the logarithmic factor needed in Corollary 7?*

Note that the statement of Corollary 7 is in regard to topological graphs in general. One may get a different answer for the restricted set of *geometric graphs*, that is, graphs with straight line segments as edges.

The geometric consequences use Theorem 1 in the special case when $n = m$, but it is interesting to give bounds in the asymmetric cases as well. We define $R(n,m)$ to be the maximum total length of m pairwise intersection reverse cyclic sequences over an alphabet of size n. With this notation Corollary 6 gives $R(n,m) = O(m\sqrt{n}\log n + n\sqrt{m})$. We collect here a few simple lower and upper bounds for $R(n,m)$.

A trivial consequence of the property that a collection of cyclic sequences are pairwise intersection reverse is that no three symbols appear together in three cyclic sequences. By the Kővári–Sós–Turán Theorem [9], we have that $R(n,m) = O(nm^{2/3} + m)$ and $R(n,m) = O(n^{2/3}m + n)$. The first bound supersedes the bound in Corollary 6 if $m \geq n^{3/2}$. The second bound supersedes the bound in Corollary 6 if $m < n^{2/3}$. So for these extremely large or small values of m Corollary 6 is not tight.

The simplest constructions of intersection reverse cyclic sequences are constructions for collections of *subsets* intersecting each other in at most two elements. No matter how we order these subsets the resulting collection of cyclic sequences is pairwise intersection reverse. A simple construction for such subsets is any collection of circles in a finite plane. Taking all points of the plane and a subset of the circles gives $R(n,m) = \Omega(m\sqrt{n})$ for $m \leq n^{3/2}$. Taking all circles and a subset of the points gives $R(n,m) = \Omega(nm^{2/3})$ for $m \geq n^{3/2}$. A collection of singleton sets gives the trivial bound $R(n,m) \geq m$, which is better than the previous bounds for $m > n^3$. Pairwise disjoint sets provide the other trivial $R(n,m) \geq n$ bound, which is better than the other bounds for $m \leq \sqrt{n}$.

The solid lines in the logarithmic scale diagram in Figure 3 shows the lower and upper bounds mentioned above. These bounds determine $R(n,m)$ up to a constant factor for $m \geq n^{3/2}$ and $m \leq n^{1/3}$ and up to a logarithmic factor for $n \leq m \leq n^{3/2}$. In any construction proving better lower bounds than the ones above, a typical pair of cyclic sequences will need to intersect in many elements, so the cyclic order becomes essential in such a construction. We present such a construction below, proving $R(n,m) = \Omega(n^{5/6}m^{1/2})$ for $n^{1/3} < m < n^{2/3}$. This bound is represented in Figure 3 by the dashed line. The area of "uncertainty" is shaded. Even with this construction, the upper and lower bounds for $R(n,m)$ are far apart for $n^{1/3} < m < n$.

Construction. The construction is based on a construction of Gy. Elekes [6] of a set of axis-aligned parabolas and a set of points with a large number of incidences. For integers $b \geq a \geq 1$ consider the subset $P = \{(i,j) : |i| \leq a, |j| \leq 3a^2b\}$ of the integer grid and consider the collection $\mathcal{C}$ of parabolas (and lines) given by $y = ux^2 + vx + w$ with integers u, v, and w satisfying $|u| \leq b$, $|v| \leq ab$ and $|w| \leq a^2b$. We have $m = |P| = (2a+1)(6a^2b+1) = \Theta(a^3b)$ and $n = |\mathcal{C}| = (2b+1)(2ab+1)(2a^2b+1) = \Theta(a^3b^3)$. Clearly, each curve in $\mathcal{C}$ contains a point in P for each possible x coordinate, a total of $2a+1$ points. For each $p \in P$ we define the linearly ordered list B_p of all the curves in $\mathcal{C}$ passing through p. We order the list B_p according to the slopes of the curves at p (breaking ties arbitrarily). As a result we get m linearly ordered lists of subsets of the set of n symbols. Since axis-aligned parabolas form a collection of pseudo-parabolas – any pair intersects at most twice (and tangent parabolas have no further points in common) – it is easy to verify that these lists are intersection reverse. Their total length is the number of incidences between P and $\mathcal{C}$, which is $\Theta(a^4b^3) = \Theta(n^{5/6}m^{1/2})$.

Problem 2. *Is it possible to find $n^{2/3}$ pairwise intersection reverse cyclic sequences over an alphabet of size n such that their total lengths sum to significantly more than $n^{7/6}$?*

Note that for $m = n^{2/3}$ both constructions give cyclic sequences with total size $\Theta(n^{7/6})$. One of the constructions is based on finite geometry, the other on Euclidean geometry. It seems to be hard to combine these constructions for a better result. The upper bound (provided both by Corollary 6 and the Kővári–Sós–Turán Theorem [9]) is $O(n^{4/3})$.

As Figure 3 shows, it is unclear as to whether the $n\sqrt{m}$ term in Corollary 6 gives a tight bound for $R(n,m)$ in any range. We claim that its appearance is meaningful, however. The total length of the sequences needs to be above this threshold in order for a typical pair of symbols to appear together in many cyclic sequences – a property which is necessary in our estimate that not many more different than same pairs exist. If a typical pair of symbols appears together in only two cyclic sequences, it is possible that they only contribute different pairs. This happens in the above construction as well; since we construct linearly ordered (rather than cyclic) sequences that are pairwise intersection reverse, no "same pair" ever appears.

One can ask the same extremal question about linearly ordered sequences. Let $Q(n,m)$ stand for the maximum total length of m pairwise intersection reverse sequences over an n element alphabet. In this case two symbols cannot appear together in three sequences. The Kővári–Sós–Turán Theorem [9] therefore gives the bounds $Q(n,m) = O(mn^{2/3} + n)$ and $Q(m,n) = O(n\sqrt{m} + m)$. For $m \le n/\log^2 n$ or $m \ge n^3$ we get the same upper bounds that we did for $R(n,m)$. The upper bound for intermediate values of m is shown by the dotted line in Figure 3. One gets simple construction of intersection reverse sequences by considering set systems with pairwise intersection limited to singletons. Just as we noted in the case of cyclic sequences, this property ensures that the sequences are pairwise intersection reverse independent of the linear order chosen. The standard construction for such set systems is the set of lines in a finite plane, yielding $Q(n,m) = \Omega(n\sqrt{m})$ for $m \ge n$ and $Q(n,m) = \Omega(m\sqrt{n})$ for $m \le n$. The bounds $Q(n,m) \ge n$ and $Q(n,m) \ge m$ are trivial (just as before). These bounds determine $Q(n,m)$ up to a constant factor for $m \le n^{1/3}$ and $m \ge n$. Notice that the construction using parabolas in the plane yield pairwise intersection reverse linearly ordered sequences and so we have $Q(n,m) = \Omega(n^{5/6}m^{1/2})$ for $n^{1/3} \le m \le n^{2/3}$. Surprisingly, the "area of uncertainty" for $Q(n,m)$ is exactly the same parallelogram as it is for $R(n,m)$. Only when $n < m < n^3$ do the bounds for $Q(n,m)$ and $R(n,m)$ diverge. We do not know if allowing for cyclic sequences can yield longer intersection reverse collections when $m < n$.

Problem 3. *Does $R(n,m) = O(Q(n,m))$ hold for $m < n$?*

As far as pseudo-circles are concerned, our result is conjectured to be far from optimal. The best known construction is a set of n pseudo-circles that needs $\Omega(n^{4/3})$ cuts before it becomes a collection of pseudo-segments.

Problem 4. *What is the tight bound for the number of non-overlapping lenses in an arrangement of n pseudo-circles?*

As noted in Section 3, the results in this paper generalize previous results in the respect that the curves no longer need to be x-monotone. However, there are certain extensions that can no longer be achieved. Chan [4] proved an *intersection-sensitive* bound, that is, a bound which is stated as a function of

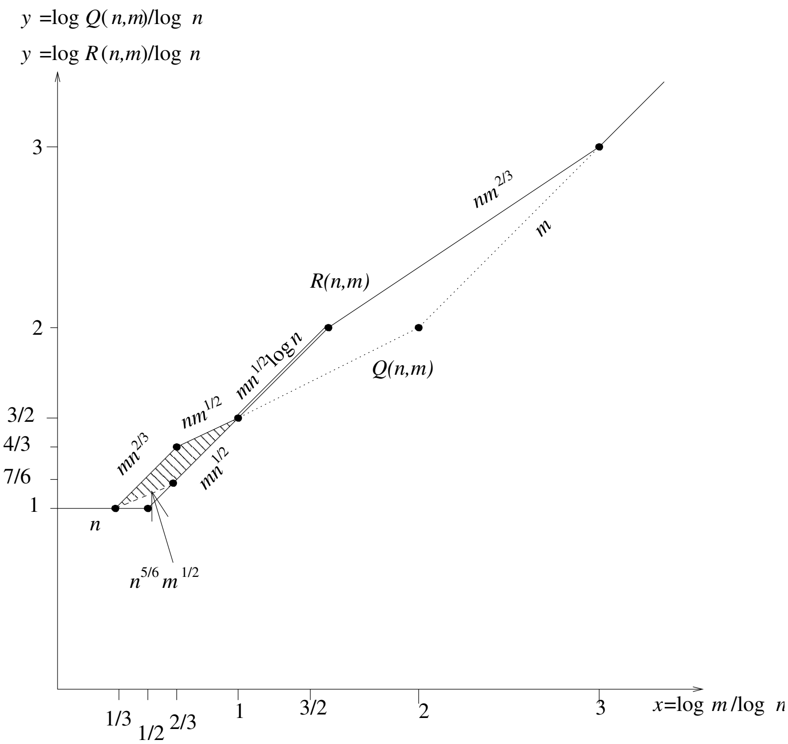

Figure 3: Bounds and area of uncertainty for $R(n, m)$ and $Q(n, m)$.

the total number of intersections. Previous papers [2, 4] are able to give such bounds for collections of x-monotone curves, but the methods break down when x-monotonicity is dropped.

Problem 5. *Find an intersection-sensitive extension to Corollary 11.*

5 Acknowledgements

We would like to thank János Pach for introducing us to this problem and for directing our attention to the many consequences that were shown in Section 3.

References

[1] P. K. Agarwal, B. Aronov, and M. Sharir, *On the complexity of many faces in arrangements of pseudo-segments and of circles*, Discrete and Computational Geometry – The Goodman-Pollack Festschrift (Heidelberg) (B. Aronov, et al., ed.), Algorithms Comb., Springer-Verlag, 2003, pp. 1–24.

[2] P. K. Agarwal, E. Nevo, J. Pach, R. Pinchasi, M. Sharir, and S. Smorodinsky, *Lenses in arrangements of pseudo-circles and their applications*, J. ACM **51** (2004), no. 2, 139–186.

[3] B. Aronov and M. Sharir, *Cutting circles into pseudo-segments and improved bounds for incidences*, Discrete Comput. Geom. **28** (2002), no. 4, 475–490.

[4] T. M. Chan, *On levels in arrangements of curves*, Discrete and Comput. Geom. **29** (2000), 375–393.

[5] ———, *On levels in arrangements of curves, II: a simple inequality and its consequence*, Proc. 44th IEEE Symposium on Foundations of Computer Science (FOCS), 2003, pp. 544–550.

[6] Gy. Elekes, *Sums versus products in number theory, algebra and Erdős geometry*, Paul Erdős and his mathematics II (G. Halász, et al., ed.), Based on the conference, Budapest, Hungary, July 4–11, 1999, no. 11, Bolyai Soc. Math. Stud., Springer, 2002, pp. 241–290.

[7] P. Erdős, *Some recent progress on extremal problems in graph theory*, Congr. Numer. **14** (1975), 3–14.

[8] T. Kaiser, *Transversals of d-intervals*, Discrete Comput. Geom. **18** (1997), no. 2, 195–203.

[9] T. Kővári, V. T. Sós, and P. Turán, *On a problem of K. Zarankiewicz*, Colloq. Math. **3** (1954), 50–57.

[10] L. Lovász, *On the ratio of optimal integral and fractional covers*, Discrete Math. **13** (1975), 383–390.

[11] R. Pinchasi and R. Radoičić, *On the number of edges in geometric graphs with no self-intersecting cycle of length 4*, Towards a Theory of Geometric Graphs (Providence, RI) (J. Pach, ed.), Contemp. Math., no. 342, Amer. Math. Soc., 2004.

[12] H. Tamaki and T. Tokuyama, *How to cut pseudo-parabolas into pseudosegments*, Discrete Comput. Geom. **19** (1998), no. 2, 265–290.

[13] G. Tardos, *Transversals of 2-intervals, a topological approach*, Combinatorica **15** (1995), no. 1, 123–134.