

REGULAR SYSTEMS OF PATHS AND FAMILIES OF CONVEX SETS IN CONVEX POSITION

MICHAEL GENE DOBBINS, ANDREAS F. HOLMSEN, AND ALFREDO HUBARD

ABSTRACT. In this paper we show that every sufficiently large family of convex bodies in the plane has a large subfamily in convex position provided that the number of common tangents of each pair of bodies is bounded and every subfamily of size five is in convex position. (If each pair of bodies has at most two common tangents it is enough to assume that every triple is in convex position, and likewise, if each pair of bodies has at most four common tangents it is enough to assume that every quadruple is in convex position.) This confirms a conjecture of Pach and Tóth and generalizes a theorem of Bisztriczky and Fejes Tóth. Our results on families of convex bodies are consequences of more general Ramsey-type results about the crossing patterns of systems of graphs of continuous functions $f : [0, 1] \rightarrow \mathbb{R}$. On our way towards proving the Pach-Tóth conjecture we obtain a combinatorial characterization of such systems of graphs in which all subsystems of equal size induce equivalent crossing patterns. These highly organized structures are what we call *regular systems of paths*, and they are natural generalizations of the notions of cups and caps from the famous theorem of Erdős and Szekeres. The characterization of regular systems is combinatorial and introduces some auxiliary structures which may be of independent interest.

1. INTRODUCTION

Loosely speaking, Ramsey theory can be summarized by the statement “any sufficiently large structure must contain a highly organized substructure”. Typically, solving a Ramsey-type problem consists of two parts. Firstly, the existence of a “highly organized substructure” is established; secondly a quantitative estimate of “sufficiently large” is given. This paper focuses on the first of the two problems. Our results rely on a crude application of Ramsey’s theorem, and therefore we expect our bounds to be quantitatively meaningless. On the other hand, deciphering what “highly organized” means in the context of families of convex bodies turned out to be the heart of the problem.

1.1. The Erdős–Szekeres theorem. Quantitative aspects of Ramsey theory were first considered in a foundational paper by Erdős and Szekeres. They were led to Ramsey’s theorem on their way to showing the following beautiful result [7, 8].

Received by the editors June 28, 2013 and, in revised form, February 28, 2014.

2010 *Mathematics Subject Classification.* Primary 05D10, 52C10; Secondary 52C30.

The first author was supported by NRF grant 2011-0030044 funded by the government of South Korea (SRC-GAIA) and BK21.

The second author was supported by Basic Science Research Program through the National Research Foundation of Korea funded by the Ministry of Education, Science and Technology (NRF-2010-0021048).

The third author was supported by Fondation Sciences Mathématiques de Paris and would like to thank KAIST for their hospitality and support during his visit.

Theorem (Erdős–Szekeres, 1935). *For every positive integer n there exists a minimal positive integer $f(n)$ such that the following holds: Any set of at least $f(n)$ points in the plane such that no three are collinear contains n points which are in convex position.*

We say that a set, or family of sets, is in *convex position* if no element, or member, is contained in the convex hull of the others. Determining the precise growth of $f(n)$ is one of the longest-standing open problems in combinatorial geometry, and as such it has generated a considerable amount of research. For every integer $n \geq 3$, Erdős and Szekeres [8] have constructed a set of 2^{n-2} points with no three collinear points, which does not contain the vertices of any convex n -gon. This example shows that the Erdős–Szekeres function $f(n)$ is strictly greater than 2^{n-2} . (See [14] for a simple description of this construction.) For $n \leq 6$ it is known that $f(n) = 2^{n-2} + 1$, and it is conjectured that this equality holds for all n [7, 8, 17]. For $n > 6$ the best known upper bound, due to Tóth and Valtr [18], is $f(n) \leq \binom{2n-5}{n-2} + 1 \sim \frac{4^n}{\sqrt{n}}$. Asymptotically this is the same as the bound given by Erdős and Szekeres in their seminal paper. For more information about the Erdős–Szekeres problem, its generalizations and related results, the reader should consult the surveys [1, 14].

1.2. Families of convex bodies. Bisztriczky and Fejes Tóth [2] extended the Erdős–Szekeres theorem to families of convex bodies in the plane. (Here a convex body means a compact convex set.)

Theorem (Bisztriczky–Fejes Tóth, 1989). *For every positive integer n there exists a minimal positive integer $h_0(n)$ such that the following holds: Any family of at least $h_0(n)$ pairwise disjoint convex bodies in the plane such that any three are in convex position contains n members which are in convex position.*

The case when the convex bodies are points shows that $f(n) \leq h_0(n)$, but bounding $h_0(n)$ is considerably more difficult. Nevertheless, Bisztriczky and Fejes Tóth conjectured that the two functions are equal, that is, $f(n) = h_0(n)$, which is known to hold for all $n \leq 6$ (see [5]). The original upper bound on $h_0(n)$ was reduced to $16^n/n$ by Pach and Tóth in [15], and in a subsequent paper [16] the disjointness condition was relaxed. They showed that it is sufficient to assume that each pair of bodies is *non-crossing*, which means that each pair of bodies has precisely two *common supporting tangents*.¹ Let $h_1(n)$ denote the corresponding Erdős–Szekeres function for non-crossing bodies, which obviously satisfies $h_0(n) \leq h_1(n)$. The upper bound of $h_1(n)$ was reduced in [9, 13], while the most substantial improvement was given by the present authors in [5], where it was shown that $h_1(n) \leq \binom{2n-5}{n-2} + 1$. Notice that this is the same bound as Tóth and Valtr’s bound for the original Erdős–Szekeres function.

1.3. The Pach–Tóth conjecture. Pach and Tóth conjectured in [16] that the non-crossing condition could be further relaxed. However, they constructed an infinite family of segments in which every three members are in convex position,

¹There are several other definitions of non-crossing convex bodies, but our definition has some technical advantages and, more importantly, does not cause any loss of generality.

but no four are, indicating that additional assumptions are needed. (See [16] for details.) Here we confirm the conjecture of Pach and Tóth.²

Theorem 1.1. *For all integers $n > k \geq 1$, there exists a minimal positive integer $h_k(n)$ such that the following holds: Any family of at least $h_k(n)$ convex bodies in the plane such that any two have at most $2k$ common supporting tangents and any m_k are in convex position contains n members which are in convex position, where $m_1 = 3$, $m_2 = 4$, and $m_k = 5$ for all $k \geq 3$.*

Surprisingly, m_k does not grow as k tends to infinity. The family of segments given by Pach and Tóth shows that the bound for m_2 cannot be reduced, and it will follow as a simple consequence of our analysis that for $k \geq 3$ the bound for m_k cannot be reduced. In particular, there exist arbitrarily large families of convex bodies such that

- any two members have precisely *six* common tangents,
- any four members are in convex position, and
- no five members are in convex position.

The construction will be given in section 7.

1.4. Systems of paths. A *path* is the graph of a continuous function $f : [0, 1] \rightarrow \mathbb{R}$, drawn on the vertical strip $[0, 1] \times \mathbb{R}$.³

Definition 1.2. A *system of paths* is a finite collection of at least three paths which satisfy the following conditions:

- The paths have distinct endpoints.
- The intersection of any pair of paths is finite.
- No two paths are tangent; paths cross at every point where they intersect.
- The intersection of any three paths is empty.

For brevity, we refer to systems of paths simply as *systems*. We say a system is *k-crossing* when each pair of paths crosses *at least 1 time and at most k times*. The *size* of a system S is the number of paths in the system and is denoted by $|S|$. By taking a subset of at least three paths of a system we obtain a *subsystem*. A path belongs to the *upper envelope* of a system if at some point in $[0, 1] \times \mathbb{R}$ it appears *above* every other path of the system. The lower envelope is defined similarly. The system is called *upper convex* if all its paths appear on the upper envelope, and *lower convex* if all its paths appear on the lower envelope.

Remark 1.3. A 1-crossing system can be viewed as a *simple* pseudoline arrangement with a distinguished horizontal direction, and in fact every simple pseudoline arrangement can be represented in this way by a “wiring diagram” (see section 5.1 of [11] or section 6.3 of [4]).

Our proof of the conjecture of Pach and Tóth applies a combinatorial analysis of *k-crossing* systems to a correspondence between upper convex systems and subfamilies in convex position. The main result of this paper provides an answer to

²Their original formulation of generalizing the non-crossing condition is by bounding the number of common boundary points among any two bodies, but it is easily seen that this is implied by bounding the number of common supporting tangents.

³What we call a path is often referred to as an *x-monotone* curve, and our terminology is used mainly for the sake of brevity.

the following Ramsey-type question: *Does every sufficiently large k -crossing system contain a large subsystem which is upper convex (or lower convex)?*

Theorem 1.4. *For any integers $k \geq 1$ and $n \geq 3$, there exists a minimal positive integer $C_k(n)$ such that the following holds. Every k -crossing system S of size at least $C_k(n)$ contains a subsystem of size n which is upper convex or lower convex.*

The case $k = 1$ is a dual version of the well-known *cups-caps* theorem of Erdős and Szekeres [7]. In this case the precise value of $C_1(n)$ is known and equals $\binom{2n-4}{n-2} + 1$. It is not hard to see that $C_k(3) = 3$ for all k , and clearly $C_k(n) \leq C_{k+1}(n)$, but as far as we know equality may hold for all k . Our bound on $C_k(n)$ is in terms of certain Ramsey numbers, and we suspect it to be very far from the truth.

Unfortunately, Theorem 1.4 does not seem to be applicable towards the conjecture of Pach and Tóth. For this we need a “one-sided” version guaranteeing the existence of a large *upper convex* subsystem, which requires the hypothesis to be strengthened accordingly.

Theorem 1.5. *For any integers $k \geq 1$ and $n \geq 3$, there exists a minimal positive integer $U_k(n)$ such that the following holds. For any k -crossing system S of size at least $U_k(n)$:*

- (1) *If $k \leq 2$ and every subsystem of size 3 is upper convex, then S contains an upper convex subsystem of size n .*
- (2) *If $k \leq 4$ and every subsystem of size 4 is upper convex, then S contains an upper convex subsystem of size n .*
- (3) *If $k \geq 5$ and every subsystem of size 5 is upper convex, then S contains an upper convex subsystem of size n .*

Our bound on $U_k(n)$ is the same as on $C_k(n)$, and again we suspect it to be far from the truth.

We are in position to prove the conjecture of Pach and Tóth.

Proof of Theorem 1.1. For a convex body K , its support function $g_K: \mathbb{S}^1 \rightarrow \mathbb{R}^1$ is defined as

$$g_K(\theta) := \max_{p \in K} \langle \theta, p \rangle.$$

(Here $\langle \cdot, \cdot \rangle$ denotes the usual Euclidean inner product.) Using this map we associate the body K with a dual support path K^* given by

$$K^* = \{(t, g_K(2\pi t)) : 0 \leq t \leq 1\}.$$

In this way a family F of convex bodies can be associated with a dual system of paths F^* , where the common supporting tangents of pairs of bodies of F are in bijective correspondence with the intersection points between pairs of dual curves of F^* . In general, F^* may not be a system, but the members of F may be perturbed in such a way that there are no tangential intersections and no triple intersections among the paths in F^* . By standard compactness arguments, this can be done without increasing the number of common supporting tangents and without changing which subfamilies are in convex position. The paths of F^* may be assumed to have distinct endpoints as well. Since any two members of F are in convex position, they must have at least two common supporting tangents. Consequently, since each pair of members of F have at most $2k$ common supporting tangents, then each pair of paths in F^* crosses at least twice and at most $2k$ times. The key observation

is that a subfamily of F is in convex position if and only if the corresponding subsystem of F^* is upper convex. Consequently, $h_k(n) \leq U_{2k}(n)$. \square

1.5. Outline of the paper. The proofs of Theorems 1.4 and 1.5 are purely combinatorial and essentially boil down to a simple parity argument: If path i starts below path j , then every time path i appears above path j , they must have crossed an odd number of times. This is of course just the final punchline, and a large part of this paper is devoted to developing the appropriate combinatorial machinery to make this formal.

In section 2 we translate the problem into combinatorial terms. The most crucial notion is that of the *local sequences* of a system. This is an encoding of the “crossing pattern” of the system, and it records, for each path, the order in which it meets the other paths. This encoding allows us to formally define what we mean by a “highly organized substructure” in a system of paths, which we refer to as *regular systems*. It is a simple consequence of Ramsey’s theorem that every sufficiently large system of paths contains a large regular subsystem. The main technical results of this paper give detailed descriptions of the envelopes of a regular system. This is the content of Theorems 2.15 and 2.16, and it is easily seen that these results imply Theorems 1.4 and 1.5.

The difficulty that arises when trying to analyze regular systems is witnessed by the fact that the number of distinct regular k -crossing systems grows very rapidly with k . The proofs of Theorems 2.15 and 2.16 span sections 3 – 6 and actually contain a complete characterization of all regular systems. Such a characterization is obtained by reformulating the notion of a regular system in terms of certain properties concerning *sequences* on an *ordered alphabet*. The discussion is organized by dividing these properties into combinatorial and geometric ones.

In section 3 we treat the combinatorial properties. Here we introduce the notion of a *tableau* which is a sequence of sequences on an ordered alphabet. The connection to our problem is that the local sequences of a system give rise to a tableau (but we do not require the converse to be true). We further define *regular tableaux* which are a combinatorial abstraction of regular systems. The main result of this section is the bijective correspondence established in Corollary 3.14 which characterizes all regular tableaux. Our results in this section are completely elementary and somewhat technical. The notion of regular tableaux seems natural, but we are unaware of connections with previously studied structures.

In section 4 we treat the geometric properties. That is, we give conditions for when a tableau corresponds to the local sequences of a system of paths. These results follow standard arguments, most of which were introduced in the study of arrangements of lines and pseudolines.

In section 5 we are ready to characterize the regular systems. This essentially amounts to describing the intersection of the combinatorial properties of section 3 and the geometric properties of section 4. As a consequence we establish the basic structure of the local sequences of any regular system.

In section 6 we apply the characterization of regular systems with the aforementioned parity argument to describe the upper and lower envelopes.

In section 7 we conclude with some final remarks and open problems.

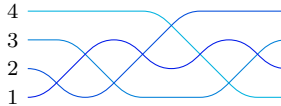
All the necessary notions are elementary and will be formally introduced along the way. As usual, the set of natural numbers is denoted by \mathbb{N} , and the finite set $\{1, 2, \dots, n\}$ is denoted by $[n]$.

2. COMBINATORIAL PROPERTIES OF SYSTEMS OF PATHS

2.1. **Local sequences.** Let S be a system and let $A \subset \mathbb{N}$ with $|A| = |S|$. Label the paths of S by the elements of A according to the order of their left endpoints from bottom to top. We say that S is *labeled* by A . Throughout the rest of this paper we will always assume that a system is labeled by an increasing sequence of positive integers.

Definition 2.1. Let S be a system labeled by $A \subset \mathbb{N}$. The *local sequence* of path i is the sequence on $A \setminus \{i\}$ which records the order in which path i intersects the other paths of the system as it is traversed from left to right.

Example 2.2. The figure below shows a system of paths labeled by $[4]$:



This gives us the corresponding local sequences:

- Path 4 : 2 1 3
- Path 3 : 1 2 4 1
- Path 2 : 1 3 1 4
- Path 1 : 2 3 2 4 3

For a finite set $A \subset \mathbb{N}$ and $i \in A$, let $A_i^- = \{j \in A : j < i\}$ and let $A_i^+ = \{j \in A : j > i\}$. The following lemma implies that the upper and lower envelopes of a system can be determined from the local sequences of its paths.

Lemma 2.3. *Let S be a system labeled by A . Path i appears on the upper envelope of S if and only if there is an initial string of its local sequence which contains every element of A_i^+ an odd number of times and every element of A_i^- an even number of times. (The same holds for the lower envelope by reversing the roles of A_i^+ and A_i^- .)*

Proof. Suppose $i < j$. This means that path i starts below path j . Therefore, every time path i appears above path j , the two paths should have crossed an odd number of times. Similarly, every time path j appears above path i , they should have crossed an even number of times. □

2.2. **Signatures of systems of size 3.** In our proofs of Theorems 1.4 and 1.5 the systems on three paths play an important role. In this case the crossings of the system are linearly ordered. (For systems of size greater than 3 we generally only have a partial ordering of the crossings.) Here we introduce a combinatorial signature which records this ordering.

Definition 2.4. Let S be a system labeled by $\{i_1, i_2, i_3\} \subset \mathbb{N}$ where $i_1 < i_2 < i_3$. The *signature* of S is the word $\sigma = \sigma(S)$ on the alphabet $\{x, y, z\}$ which records the linear ordering of the crossings of S by the rules

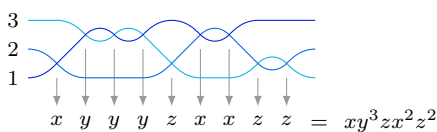
$$\{i_1, i_2\}\text{-crossing} \mapsto x \quad , \quad \{i_1, i_3\}\text{-crossing} \mapsto y \quad , \quad \{i_2, i_3\}\text{-crossing} \mapsto z.$$

Remark 2.5. We consider the alphabet $\{x, y, z\}$ to be ordered $x \prec y \prec z$. This will be crucial later on. Notice that this just corresponds to the lexicographical ordering

of the pairs (i_1, i_2) , (i_1, i_3) , (i_2, i_3) , but introducing letters x, y, z simplifies the notation.

Remark 2.6. The signature of a system of size 3 is a compact way of encoding the local sequences of the system. Let S be a system labeled by [3] with signature σ . Let $\sigma_{\{x,y\}}$ denote the word obtained by deleting the letter z from σ . If we replace each x by 2 and each y by 3 in $\sigma_{\{x,y\}}$, then we obtain the local sequence of path 1. This follows from the definition of the signature. Similarly, let $\sigma_{\{x,z\}}$ and $\sigma_{\{y,z\}}$ be the words obtained by deleting the letters y and x from σ , respectively. Replacing each x by 1 and each z by 3 in $\sigma_{\{x,z\}}$ gives us the local sequence of path 2, and replacing each y by 1 and each z by 2 in $\sigma_{\{y,z\}}$ gives us the local sequence of path 3.

Example 2.7. The system depicted in the figure below has signature $\sigma = xy^3zx^2z^2$:



The words $\sigma_{\{x,y\}}$, $\sigma_{\{x,z\}}$, and $\sigma_{\{y,z\}}$ give us the following local sequences:

$$\begin{aligned} \sigma_{\{y,z\}} &= y^3z^3 \longrightarrow \text{Path 3 : } 1 \ 1 \ 1 \ 2 \ 2 \ 2 \\ \sigma_{\{x,z\}} &= xzx^2z^2 \longrightarrow \text{Path 2 : } 1 \ 3 \ 1 \ 1 \ 3 \ 3 \\ \sigma_{\{x,y\}} &= xy^3x^2 \longrightarrow \text{Path 1 : } 2 \ 3 \ 3 \ 3 \ 2 \ 2 \end{aligned}$$

Proposition 2.8. Let S be a system of size 3 with signature σ . Suppose $u, v, w \in \{x, y, z\}$ where $u \neq v$ and $w \neq v$. The following hold.

- (1) For $\sigma = u^pv \dots$ we have $u \in \{x, z\}$ and $v = y \iff p$ is odd.
- (2) For $\sigma = \dots uv^pw \dots$ we have $u = w \iff p$ is even.

Proof. The first claim of (1) is obvious: The first crossing which occurs must involve path i_2 , since a crossing always involves the path which is currently in the middle, and i_2 starts in the middle. Path i_2 is in the middle after the first p crossings if and only if p is even; hence, the $(p+1)$ 'st crossing is an $\{i_1, i_3\}$ -crossing if and only if p is odd. For part (2), suppose the u corresponds to an $\{i, j\}$ -crossing, where j is the path which is in the middle *after* this crossing occurs. This means that the next p crossings which correspond to v^p involve path j . The w corresponds to an $\{i, k\}$ -crossing, and we suppose k is the path which is in the middle *before* this crossing occurs. Clearly, $k = j$ if and only if p is even. □

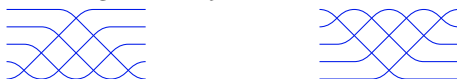
Remark 2.9. It is not hard to verify that Proposition 2.8 characterizes the set of signatures. That is, any word on the alphabet $\{x, y, z\}$ which satisfies the conditions of Proposition 2.8 corresponds to the signature of a system of size 3. We leave the proof to the reader.

2.3. Regular systems.

Definition 2.10. A system S is *regular* if it has size at least 4 and the signatures $\sigma(T) = \sigma(T') \neq \emptyset$ for all subsystems T and T' of size 3. The unique signature of the subsystems of size 3 is called the *signature* of the regular system.

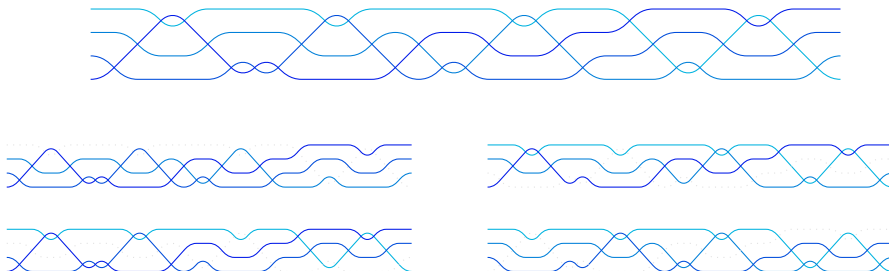
Remark 2.11. Note that in a regular system each pair of paths must cross the same number of times and that any subsystem of size at least 4 is also regular.

Example 2.12. There are precisely two distinct signatures of 1-crossing systems of size 3, and for any $n \geq 4$ there exist regular systems of size n with these signatures. The system below on the left is regular with signature xyz , while the system below on the right is regular with signature zyx .



One can think of these systems as the dual to what Erdős and Szekeres call cups and caps [7], and their result states that any 1-crossing system of size $\binom{2n-4}{n-2} + 1$ contains a regular subsystem of size n .

Example 2.13. The figure below shows a system S of size 4 together with its 4 subsystems of size 3. It is easily seen that each subsystem of size 3 has signature $\sigma = xy^2x^3z^2xyz^2y^2z$. Therefore S is a regular system.



Example 2.14. The figure below shows a system of size 3 with signature $\sigma = xy^2x^4yz^4y^2z$.



As we will see, there exists no regular system with signature σ . This particular example will be revisited in Example 5.9.

2.4. The existence of large regular systems. For any fixed integer $k \geq 1$, it follows from Ramsey's theorem that every sufficiently large k -crossing system contains a large subsystem which is regular. The argument goes as follows. Let $R_3(n; M)$ denote the *symmetric Ramsey number* for M -partitions of the edge set of the complete 3-uniform hypergraph. In other words, for $N \geq R_3(n; M)$, every partition of the triples of $[N]$ into at most M classes, there exists a subset $A \subset [N]$, with $|A| \geq n$, such that every triple of A belongs to the same class. For every $k \geq 1$, there is a finite number of signatures of k -crossing systems of size 3. Let M_k denote this number. Let S be a k -crossing system of size $N \geq R_3(n; M_k)$. If we partition the subsystems of size 3 according to their signature, Ramsey's theorem implies that S contains a regular system of size n .

2.5. Envelopes of regular systems. In a regular system, a path will typically cross all paths above or below it going directly to an envelope, and if this happens for one path, the same happens for all paths. This behavior can be seen in Example 2.13 and results in either very few or all paths appearing on the upper or lower envelopes. This is made precise in Theorems 2.15 and 2.16, below, and since every

sufficiently large k -crossing system contains a large subsystem which is regular (by the argument in section 2.4), these statements easily imply Theorems 1.4 and 1.5.

Theorem 2.15. *Every regular system is upper convex or lower convex.*

Proof of Theorem 1.4. By the argument in section 2.4, if S is a k -crossing system of size at least $R_3(n; M_k)$, then S contains a subsystem S' of size n which is regular. By Theorem 2.15, S' is upper convex or lower convex. Therefore $C_k(n) \leq R_3(n; M_k)$. □

Theorem 2.16. *Let S be a regular system.*

- (1) *If S is 2-crossing, then S is upper convex or only 2 paths appear on the upper envelope of S .*
- (2) *If S is 4-crossing, then S is upper convex or at most 3 paths appear on the upper envelope of S .*
- (3) *If S is k -crossing for $k > 4$, then S is upper convex or at most 4 distinct paths appear on the upper envelope of S .*

Proof of Theorem 1.5. If S is a k -crossing system of size at least $R_3(n; M_k)$, then S contains a regular subsystem S' of size n . For (1) of Theorem 1.5 suppose S is 2-crossing. By (1) of Theorem 2.16, S' is upper convex or only 2 paths appear on the upper envelope of S' . The latter case is impossible: If only 2 paths appear on the upper envelope, then S' contains a subsystem of size 3 which is not upper convex (take the two paths from the upper envelope together with any other path of S'). This contradicts the hypothesis. Similarly, (2) and (3) of Theorem 1.5 are implied by (2) and (3) of Theorem 2.16, respectively. Therefore $U_k(n) \leq R_3(n; M_k)$. □

3. SEQUENCES AND TABLEAUX

3.1. Regular sequences. For a totally ordered alphabet A , a word is a finite sequence of letters in A , and a language is a set of words. We will generally refer to words on $A \subset \mathbb{N}$ as sequences. Let A^* denote the language of all words on A , and for a language L , let L^* denote the language of all words formed by concatenating words in L . For words ω and ψ in A^* , let $\omega \cdot \psi$ denote the concatenation of ω and ψ , let $|\omega|$ denote the length of ω , and let $[\omega]$ denote the set of distinct letters appearing in ω . For $a \in A$, let a^i be the word $a \cdot \dots \cdot a$ with $|a^i| = i$. For a subset $X \subset A$, let ω_X denote the subword of ω consisting of letters in X . We call this the *restriction* of ω to X . Define a map $N: A^* \rightarrow \mathbb{N}^*$ as follows. For $\omega \in A^*$ with $[\omega] = \{a_1, \dots, a_k\}$ where $a_i < a_{i+1}$, let $N(\omega)$ denote the sequence in \mathbb{N}^* obtained by the map $a_i \mapsto i$. For instance, $N(4, 2, 6, 6, 7) = (2, 1, 3, 3, 4)$. Two words ω and ω' are *order equivalent* if $N(\omega) = N(\omega')$, in which case we write $\omega \sim \omega'$. Note that order equivalence is an equivalence relation. The map N and the notion of order equivalence naturally extend to words which are not defined on the same alphabets. For instance, if A is the Latin alphabet, then $N(\mathbf{h e l l o}) = (2, 1, 3, 3, 4)$, so $(4, 2, 6, 6, 7) \sim \mathbf{h e l l o}$. Let $\omega(a, i)$ denote the prefix of ω ending with the i 'th occurrence of the letter a . In particular, $|\omega(a, i)|$ is the position of the i 'th occurrence of a . For instance, if $\omega = \mathbf{S z e k e r e s}$ then $\omega(\mathbf{e}, 2) = \mathbf{S z e k e}$ and $|\omega(\mathbf{e}, 2)| = 5$.

Definition 3.1. Let $A \subset \mathbb{N}$ and $\omega \in A^*$. We say ω is a *regular* sequence on A if the following hold.

- A consists of at least 3 elements.
- $\omega_X \sim \omega_Y$ for all subsets $X, Y \subset A$ of size 2.

Example 3.2. The sequence $\omega = (1, 2, 3, 3, 3, 2, 2, 1, 1)$ is regular on $[3]$. The restrictions are:

$$\omega_{\{1,2\}} = (1, 2, 2, 2, 1, 1), \quad \omega_{\{1,3\}} = (1, 3, 3, 3, 1, 1), \quad \omega_{\{2,3\}} = (2, 3, 3, 3, 2, 2).$$

Example 3.3. The sequence $\omega = (1, 2, 3, 1, 3, 2)$ is not regular since the restrictions $\omega_{\{1,2\}} = (1, 2, 1, 2)$ and $\omega_{\{2,3\}} = (2, 3, 3, 2)$ are not order equivalent.

A regular sequence on A is uniquely determined by A and its restriction to any two elements.

Lemma 3.4. *If ψ and ω are regular sequences on $[n]$ and $\psi_{\{1,2\}} = \omega_{\{1,2\}}$, then $\psi = \omega$.*

Proof. Suppose the lemma fails. Let the first entry where ψ and ω differ be the p 'th occurrence of i in ψ and the q 'th occurrence of j in ω . This implies that $\psi_{\{i,j\}}$ and $\omega_{\{i,j\}}$ coincide for the first $p + q - 2$ entries, but at the $p + q - 1$ 'th entry they differ, which contradicts the fact that they are both order equivalent to $\omega_{\{1,2\}}$. \square

We define a language characterizing the restrictions of regular sequences to subsets of size 2. Let $\{a, b\}$ be an ordered alphabet with $a < b$. A *balanced block* on the alphabet $\{a, b\}$ is a word of the form

$$a^r b^r = \underbrace{a \cdots a}_{\times r} \underbrace{ab \cdots ab}_{\times r} \underbrace{b \cdots b}_{\times r} \quad \text{or} \quad b^r a^r = \underbrace{b \cdots b}_{\times r} \underbrace{ba \cdots ba}_{\times r} \underbrace{a \cdots a}_{\times r} \quad \text{for some } r \in \mathbb{N}.$$

Let $B_{(ab)}$ denote the language of balanced blocks on $\{a, b\}$. A word $\omega \in B_{(ab)}^*$ is called *balanced*; that is, words $\omega = \omega_1 \cdots \omega_k$ where $\omega_i = a^{r_i} b^{r_i}$ or $b^{r_i} a^{r_i}$. We define the *block sizes* of ω to be the sequence (r_1, r_2, \dots, r_k) , which we denote by $\langle \omega \rangle$. For instance, $\langle (a^2 b^2)(ba)(ab) \rangle = (2, 1, 1)$. Here is a simple way to check if a word is balanced. Starting from the first letter parsing one letter at a time, count the number of occurrences of a and the number of occurrences of b . Whenever a is succeeded by b , the number of a 's counted so far must be at least the number of b 's, and analogously when b is succeeded by a . In the end, we must have counted the same number of each letter. That is, we have the following equivalent definition.

Lemma 3.5. *A word $\omega \in \{a, b\}^*$ is balanced if and only if the following hold.*

- $|\omega_{\{a\}}| = |\omega_{\{b\}}|$.
- If $|\omega(a, i)| + 1 = |\omega(b, j + 1)|$, then $i \geq j$.
- If $|\omega(b, i)| + 1 = |\omega(a, j + 1)|$, then $i \geq j$.

Proof. Let ω satisfy the conditions of the lemma. We will show that ω is balanced by expressing ω as a concatenation of balanced blocks. We may assume by symmetry that ω begins with $a^j b \cdots$. If there are no more occurrences of a , then $\omega = a^j b^j$ is balanced. Otherwise ω begins with $a^j b^i a \cdots$, and since $|\omega(b, i)| + 1 = |\omega(a, j + 1)|$ we have $i \geq j$, so the first $2j$ letters of ω are the balanced block $\omega_1 = a^j b^j$.

Let $\omega = \omega_1 \cdot \omega_2$. Now we have $|\omega_{2\{a\}}| = |\omega_{\{a\}}| - j$ and $|\omega_2(a, i - j)| = |\omega(a, i)| - 2j$, and likewise for occurrences of the letter b in ω_2 , so ω_2 satisfies the conditions of the lemma. Therefore, by induction on the length of ω , we can express ω as a concatenation of balanced blocks.

The other direction follows immediately from the definition of balanced blocks and the fact that the conditions of the lemma are preserved by concatenation. \square

As a consequence we have the following.

Corollary 3.6. *If $\omega = \omega_1 \cdot \omega_2$ is balanced, then the following are equivalent.*

- ω_1 is balanced.
- ω_2 is balanced.
- $|\omega_{1\{a\}}| = |\omega_{1\{b\}}|$.
- $|\omega_{2\{a\}}| = |\omega_{2\{b\}}|$.

Proposition 3.7. *For any $n \geq 3$ and $\omega \in \{a, b\}^*$, there is a regular sequence ν on $[n]$ with $\nu_{\{1,2\}} \sim \omega$ if and only if ω is balanced.*

Proof. We define a function, $\varphi_n : B_{(ab)}^* \rightarrow [n]^*$, which produces a regular sequence on $[n]$ such that $\varphi_n(\omega)_{\{1,2\}} \sim \omega$. First, we define φ_n on balanced blocks as

$$\begin{aligned} \varphi_n(a^r b^r) &:= 1, \dots, \underset{\times r}{1}, \underset{\times r}{2}, \dots, \underset{\times r}{n}, \dots, \underset{\times r}{n} \\ \varphi_n(b^r a^r) &:= \underset{\times r}{n}, \dots, \underset{\times r}{n}, \dots, \underset{\times r}{2}, \dots, \underset{\times r}{2}, \underset{\times r}{1}, \dots, \underset{\times r}{1}. \end{aligned}$$

Extend φ_n to arbitrary words $\omega \in B_{(ab)}^*$ by concatenation in the following way: For $\omega = \omega_1 \cdots \omega_k$, where $\omega_i \in B_{(ab)}$, define $\varphi_n(\omega) := \varphi_n(\omega_1) \cdots \varphi_n(\omega_k)$. Note that the function φ_n is well defined. Let $\nu = \varphi_n(\omega)$. Clearly $\nu_X \sim \omega$ for any $X \subset [n]$ with $|X| = 2$, so the sequence $\nu = \varphi_n(\omega)$ is regular.

For the other direction, let ν be a regular sequence on $[n]$, and suppose $\nu_{\{1,2\}}$ is not balanced. Then, by Lemma 3.5 and by symmetry, we may assume there is $i < j - 1$ such that $|\nu(1, i)| + 1 = |\nu(2, j)|$. Since $\nu_{\{1,2\}} \sim \nu_{\{1,3\}}$, the j 'th occurrence of 3 must be between the i 'th and $i + 1$ 'th occurrence of 1, $|\nu(1, i)| < |\nu(3, j)| < |\nu(1, i + 1)|$, and since $\nu_{\{1,2\}} \sim \nu_{\{2,3\}}$, the j 'th occurrence of 3 must also be between the i 'th and $i + 1$ 'th occurrence of 2, $|\nu(2, i)| < |\nu(3, j)| < |\nu(2, i + 1)|$. But this is impossible, since $|\nu(2, i + 1)| \leq |\nu(2, j)| - 1 = |\nu(1, i)|$. Thus, $\nu_{\{1,2\}}$ is balanced. □

Corollary 3.8. *Let $n \geq 3$ be fixed. The set of regular sequences on $[n]$ is in bijective correspondence with the language $B_{(ab)}^*$.*

Proof. This follows from Lemma 3.4 and Proposition 3.7. □

Example 3.9. Consider $\omega = (ab)(a^2b^2)(ba) \in B_{(ab)}^*$. For $n = 3$ we get

$$\nu = \varphi_3(\omega) = (1, 2, 3, 1, 1, 2, 2, 3, 3, 3, 2, 1).$$

The restrictions are:

$$\nu_{\{1,2\}} = (1, 2, 1, 1, 2, 2, 2, 1), \quad \nu_{\{1,3\}} = (1, 3, 1, 1, 3, 3, 3, 1), \quad \nu_{\{2,3\}} = (2, 3, 2, 2, 3, 3, 3, 2).$$

3.2. Regular tableaux. Let $A = \{a_1, \dots, a_n\} \subset \mathbb{N}$ with $a_i < a_{i+1}$. A sequence $T = (\omega_1, \dots, \omega_n)$ where $\omega_i \in (A \setminus \{a_i\})^*$ is called a *tableau* on A . We call ω_i the i 'th row of T , and denote it by $T(i) := \omega_i$. Let A^\square be the set of all tableaux on A . For a subset $X = \{a_{i_1}, \dots, a_{i_k}\} \subset A$ with $i_j < i_{j+1}$, we define a tableau $T_X := (T(i_1)_X, \dots, T(i_k)_X) \in X^\square$. We call this the *restriction* of T to X . Define $N : A^\square \rightarrow \mathbb{N}^\square$ by $a_i \mapsto i$. The tableaux T and T' are *order equivalent* when $N(T) = N(T')$, in which case we write $T \sim T'$. Given two tableaux T_1 and T_2 on A , we define their concatenation, $T_1 \cdot T_2$, by letting $(T_1 \cdot T_2)(i) = T_1(i) \cdot T_2(i)$ for every $i \in A$.

When considering specific examples of tableaux it is convenient to use a graphical representation. We represent a tableau as left-justified rows of boxes in increasing order from bottom to top containing the letters of each row in their given order. For

instance, let $T = (\omega_1, \omega_2, \omega_3)$ be a tableau on $\{4, 7, 9\}$ with rows $\omega_1 = (7, 9, 7, 9, 9, 7)$, $\omega_2 = (9, 9, 4, 4)$, and $\omega_3 = (4, 7, 7, 4, 7)$. Its graphical representation is given below along with $N(T)$.

$$T = \begin{array}{|c|c|c|c|c|c|} \hline 4 & 7 & 7 & 4 & 7 & \\ \hline 9 & 9 & 4 & 4 & & \\ \hline 7 & 9 & 7 & 9 & 9 & 7 \\ \hline \end{array} \qquad N(T) = \begin{array}{|c|c|c|c|c|} \hline 1 & 2 & 2 & 1 & 2 \\ \hline 3 & 3 & 1 & 1 & \\ \hline 2 & 3 & 2 & 3 & 3 & 2 \\ \hline \end{array}$$

Definition 3.10. Let T be a tableau on $A \subset \mathbb{N}$. We say T is *regular* if the following hold.

- A consists of at least 4 elements.
- $T_X \sim T_Y$ for all subsets $X, Y \subset A$ of size 3.

Example 3.11. Below is a regular tableau on $[4]$.

$$T = \begin{array}{|c|c|c|} \hline 1 & 2 & 3 \\ \hline 4 & 4 & 1 & 2 \\ \hline 4 & 4 & 3 & 3 & 1 \\ \hline 4 & 4 & 3 & 3 & 2 & 2 \\ \hline \end{array}$$

Restricting to the 3-subsets of $[4]$ gives us the following tableaux:

$$T_{\{1,2,3\}} = \begin{array}{|c|c|} \hline 1 & 2 \\ \hline 3 & 3 & 1 \\ \hline 3 & 3 & 2 & 2 \\ \hline \end{array} \quad T_{\{1,2,4\}} = \begin{array}{|c|c|} \hline 1 & 2 \\ \hline 4 & 4 & 1 \\ \hline 4 & 4 & 2 & 2 \\ \hline \end{array} \quad T_{\{1,3,4\}} = \begin{array}{|c|c|} \hline 1 & 3 \\ \hline 4 & 4 & 1 \\ \hline 4 & 4 & 3 & 3 \\ \hline \end{array} \quad T_{\{2,3,4\}} = \begin{array}{|c|c|} \hline 2 & 3 \\ \hline 4 & 4 & 2 \\ \hline 4 & 4 & 3 & 3 \\ \hline \end{array}$$

A regular tableau on A is uniquely determined by A and its restriction to any three elements.

Lemma 3.12. *If T and U are regular tableaux on $[n]$ such that $T_{\{1,2,3\}} = U_{\{1,2,3\}}$, then $T = U$.*

Proof. Suppose the lemma fails. Then there is some $i \in [n]$ such that $T(i) \neq U(i)$. Consider the first entry where they differ; say the m 'th entry of $T(i)$ is j and the m 'th entry of $U(i)$ is k . As in the proof of Lemma 3.4, $T(i)_{\{j,k\}} \neq U(i)_{\{j,k\}}$, so $T_{\{i,j,k\}} \neq U_{\{i,j,k\}}$. But this is impossible, since these are both the unique tableau on $\{i, j, k\}$ which is order equivalent to $T_{\{1,2,3\}}$. \square

A sequence $(s_1, s_2, \dots, s_n) \in \mathbb{N}^*$ is a *refinement* of (t_1, t_2, \dots, t_m) when there exist integers i_j such that

$$(t_1, t_2, \dots) = (s_1 + \dots + s_{i_1}, s_{i_1+1} + \dots + s_{i_2}, \dots).$$

The *exponent sequence* of a letter a in a word ω is the sequence of sizes of consecutive occurrences of a in ω and is denoted by $\text{exp}_a(\omega)$. For instance, $\text{exp}_a(a^2bac^3a^5) = (2, 1, 5)$.

Proposition 3.13. *For any $n \geq 4$ and $U \in [3]^{\square}$, there is a regular tableau T on $[n]$ with $T_{\{1,2,3\}} = U$ if and only if the following hold.*

- (1) Rows $U(1)$ and $U(3)$ are balanced.
- (2) The block sizes of $U(1)$ are a refinement of $\text{exp}_3(U(2))$.
- (3) The block sizes of $U(3)$ are a refinement of $\text{exp}_1(U(2))$.

The proof will provide a correspondence analogous to that of Corollary 3.8. Let $\{a, b, c, d\}$ be an ordered alphabet with $a \prec b \prec c \prec d$, and let $B_{(ab,cd)} = B_{(ab)} \cup B_{(cd)}$. Roughly speaking, the first and the last rows of a regular tableau

are regular sequences corresponding to words $\gamma \in B_{(cd)}^*$ and $\alpha \in B_{(ab)}^*$, respectively. The intermediate rows are an interpolation between the first and the last rows described by interlacing the balanced blocks of α and γ .

Corollary 3.14. *Let $n \geq 4$ be fixed. The set of regular tableaux on $[n]$ is in bijective correspondence with the language $B_{(ab,cd)}^*$.*

Proof of Proposition 3.13. We define a function $\varphi_n : B_{(ab,cd)}^* \rightarrow [n]^{\square}$ such that the range of φ_n consists of all regular tableaux on $[n]$. The rows of the tableaux are given by $\varphi_{i,n} : B_{(ab,cd)}^* \rightarrow [n]^*$, that is,

$$\varphi_n(\omega) = (\varphi_{1,n}(\omega), \varphi_{2,n}(\omega), \dots, \varphi_{n,n}(\omega)).$$

We first define $\varphi_{i,n}$ on balanced blocks as

$$\begin{aligned} \varphi_{i,n}(a^r b^r) &:= (1, \dots, 1, \dots, i-1, \dots, i-1) \text{ for } i > 1, \varphi_{1,n}(a^r b^r) = \emptyset \\ \varphi_{i,n}(b^r a^r) &:= (i-1, \dots, i-1, \dots, 1, \dots, 1) \text{ for } i > 1, \varphi_{1,n}(b^r a^r) = \emptyset \\ \varphi_{i,n}(c^r d^r) &:= (i+1, \dots, i+1, \dots, n, \dots, n) \text{ for } i < n, \varphi_{n,n}(c^r d^r) = \emptyset \\ \varphi_{i,n}(d^r c^r) &:= (n, \dots, n, \dots, i+1, \dots, i+1) \text{ for } i < n, \varphi_{n,n}(d^r c^r) = \emptyset. \end{aligned}$$

Extend $\varphi_{i,n}$ to arbitrary words $\omega \in B_{(ab,cd)}^*$ by concatenation in the following way: For $\omega = \omega_1 \cdots \omega_k$ where $\omega_j \in B_{(ab,cd)}$, let $\varphi_{i,n}(\omega) := \varphi_{i,n}(\omega_1) \cdots \varphi_{i,n}(\omega_k)$. Since the functions $\varphi_{i,n}$ are injective, the function φ_n is also injective. Let $T^\omega = \varphi_n(\omega)$. To see that the tableau is regular, let $X = \{i, j, k\}$ with $1 \leq i < j < k \leq n$. Now we have $T^\omega(i)_X \sim \omega_{\{c,d\}}$, $T^\omega(j)_X \sim \omega_{\{a,c\}}$, and $T^\omega(k)_X \sim \omega_{\{a,b\}}$.

Let $U \in [3]^{\square}$ satisfy conditions (1), (2) and (3). We will define ω such that $T_{\{1,2,3\}}^\omega = U$ by interlacing the balanced blocks of words in $B_{(cd)}^*$ and $B_{(ab)}^*$ corresponding to $U(1)$ and $U(3)$ according to the refinements of the exponent sequences of $U(2)$. Let

$$U(2) = (1, \dots, 1, 3, \dots, 3, 1, \dots, 1, 3, \dots, 3, \dots)$$

$\times p_1$ $\times q_1$ $\times p_2$ $\times q_2$

where $p_1 \geq 0$ and any subsequent defined terms q_i or p_i are positive. Let $\alpha \in \{a, b\}^*$ be order equivalent to $U(3)$, and $\gamma \in \{c, d\}^*$ be order equivalent to $U(1)$. Let $\alpha = \alpha_1 \alpha_2 \cdots$ and $\gamma = \gamma_1 \gamma_2 \cdots$ such that $|\alpha_i| = 2p_i$ and $|\gamma_i| = 2q_i$, and let $\omega = \alpha_1 \gamma_1 \alpha_2 \gamma_2 \cdots$.

We first show $\omega \in B_{(ab,cd)}^*$. The block sizes of $\gamma \sim U(1)$ are a common refinement of $(q_1, q_2, \dots) = \text{exp}_3(S_2)$. This implies that the words γ_i are balanced, and likewise the α_i are balanced, so $\omega \in B_{(ab,cd)}^*$. We now show $T_{\{1,2,3\}}^\omega = U$. By construction, $\omega_{\{c,d\}} = \gamma \sim U(1)$, so $T^\omega(1)_{\{2,3\}} = U(1)$, and similarly $T^\omega(3)_{\{1,2\}} = U(3)$. Lastly, $\omega_{\{a,c\}} \sim U(2)$, since $|\alpha_{i_{\{a\}}}| = p_i$ and $|\gamma_{i_{\{c\}}}| = q_i$, so $T^\omega(2)_{\{1,3\}} = U(2)$. Thus, we have a regular tableau T^ω such that $T_{\{1,2,3\}}^\omega = U$.

For the other direction let T be a regular tableau on $[n]$, and let $U = T_{\{1,2,3\}}$.

We first show that $U(1)$ and $U(3)$ are balanced. We claim that $T(1)$ and $T(n)$ are regular sequences on $[n]_1^+$ and $[n]_n^-$, respectively. For any $X = \{1, i, j\}$ and $Y = \{1, i', j'\}$, the regularity of T implies that $T(1)_X \sim T(1)_Y$. Therefore $T(1)$ is regular on $[n]_1^+$. The same reasoning, restricting to sets $X = \{i, j, n\}$ and $Y = \{i', j', n\}$, shows $T(n)$ is regular. Now, by Proposition 3.7, $U(1) = T(1)_{\{2,3\}}$ is balanced, since $T(1)$ is regular, and likewise $U(3) = T(n)_{\{1,2\}}$ is balanced, since $T(n)$ is regular.

Now we show that the refinements hold. By definition $n \geq 4$. Since $T_{\{1,2,3\}} \sim T_{\{1,2,4\}}$, we can let

$$(t_1, t_2, \dots, t_m) = \exp_3(T(2)_{\{1,3\}}) = \exp_4(T(2)_{\{1,4\}}).$$

We decompose $T(2)_{\{3,4\}}$ into subsequences appearing between consecutive occurrences of 1 in $T(2)_{\{1,3,4\}}$; that is, let $T(2)_{\{3,4\}} = \omega_1 \cdot \omega_2 \cdots \omega_m$ such that

$$T(2)_{\{1,3,4\}} = (\dots, 1, \omega_i, 1, \dots, 1, \omega_{i+1}, 1, \dots).$$

Since $T_{\{2,3,4\}} \sim U$, $T(2)_{\{3,4\}} \sim U(1)$ is balanced, and since $|\omega_{i_{\{3\}}}| = |\omega_{i_{\{4\}}}| = t_i$ for all i , each subsequence ω_i is balanced by Corollary 3.6. Let $(s_{(i,1)}, \dots, s_{(i,k_i)})$ be the block sizes of ω_i . The block sizes of $U(1)$ are $(s_{(1,1)}, \dots, s_{(1,k_1)}, s_{(2,1)}, \dots, s_{(m,k_m)})$, and $s_{(i,1)} + \dots + s_{(i,k_i)} = t_i$. Therefore, the block sizes of $U(1)$ are a refinement of $\exp_3(U(2))$, and likewise for $U(2)$ and $\exp_1(U(2))$. \square

Example 3.15. Consider the word $(ab)(d^2c^2)(ba) \in B_{(ab,cd)}^*$. For $n = 4$, we get the following tableau:

1	2	3	3	2	1
1	2	4	4	2	1
1	4	4	3	3	1
4	4	3	3	2	2

Restricting to the 3-subsets of [4] gives us the following tableaux:

1	2	2	1
1	3	3	1
3	3	2	2

1	2	2	1
1	4	4	1
4	4	2	2

1	3	3	1
1	4	4	1
4	4	3	3

2	3	3	2
2	4	4	2
4	4	3	3

Let $T_n^\omega := \varphi_n(\omega)$ denote the regular tableau on $[n]$ corresponding to $\omega \in B_{(ab,cd)}^*$. Note that T_n^ω is well defined for all $n \geq 3$, and that T_n^ω is regular for all $n \geq 4$. Notice that $T_5^\omega(3) \sim \omega$. We point out two further consequences of the proof of Proposition 3.13.

Corollary 3.16. *If $T = T_n^\omega$ is a regular tableau and X is a 3-subset of $[n]$, then $T_X \sim T_3^\omega$.*

Corollary 3.17. *If $T_n^{\omega_1}$ and $T_n^{\omega_2}$ are regular tableaux, then $T_n^{\omega_1} \cdot T_n^{\omega_2} = T_n^{\omega_1 \cdot \omega_2}$.*

A tableau $T \in A^n$ is called *pangrammatic* when every $a_i \in A$ occurs at least once in T . In the following sections all regular tableaux are assumed to be pangrammatic. Let $W_{(ab,cd)}^*$ be the language of words $\omega \in B_{(ab,cd)}^*$ such that $\omega \notin B_{(ab)}^*$ and $\omega \notin B_{(cd)}^*$.

Proposition 3.18. *Let $n \geq 4$ be fixed. The set of pangrammatic regular tableaux on $[n]$ is in bijective correspondence with the language $W_{(ab,cd)}^*$.*

Proof. Observe from the definition of φ_n in the proof of Proposition 3.13 that for $\omega \in B_{(ab)}$, the integers occurring in T_n^ω are $\{1, \dots, n - 1\}$, and for $\omega \in B_{(cd)}$, the integers occurring in T_n^ω are $\{2, \dots, n\}$. \square

4. GEOMETRIC TABLEAUX

4.1. Tableaux and local sequences. Let S be a system labeled by $[n]$, and let ω_i denote the local sequence of path i . We may associate the local sequences of S with a tableau $T = (\omega_1, \dots, \omega_n)$ on $[n]$. We call T the tableau *associated* to S .

Clearly $T \in [n]^p$. For us, the converse direction is important. That is, when does a tableau on $[n]$ correspond to the local sequences of a system of paths?

Definition 4.1. Let $T \in [n]^p$. We say T is *geometric* if it is the associated tableau of a system.

We can decide if a tableau is geometric by a simple algorithm.

Algorithm 4.2. *The input is a tableau T on $[n]$. The output is **True** if T is geometric and **False** otherwise.*

```

T' ← T
π ← (1, ..., n).
while T' ≠ (∅, ..., ∅) do
  if There exists a lexicographically minimal pair of integers (j, k) that satisfy
  the following.
    • j and k are adjacent in π.
    • T'(j) = (k) · ωj.
    • T'(k) = (j) · ωk.
  then
    Transpose elements j and k in π.
    T'(j) ← ωj.
    T'(k) ← ωk.
  else
    Return False.
  end if
end while
Return True.
    
```

Proposition 4.3. *A tableau T on $[n]$ is geometric if and only if Algorithm 4.2 returns **True** on T .*

Proof. Consider a system S ordered by $[n]$ and let T be the associated tableau. We show that applying Algorithm 4.2 to T corresponds to a “topological sweep” of S .⁴ Define a *cut path* of S to be a continuous path contained in the strip $[0, 1] \times \mathbb{R}$ that

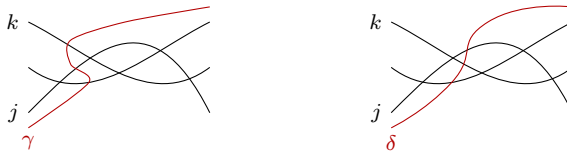
- starts on the line $\{0\} \times \mathbb{R}$ below the left endpoints of the paths of S ,
- ends on the line $\{1\} \times \mathbb{R}$ above the right endpoints of the paths of S , and
- intersects each path of S at a unique point in $(0, 1) \times \mathbb{R}$.

We do not require a cut path to be the graph of a function $f: [0, 1] \rightarrow \mathbb{R}$, but the conditions imply that it intersects every path in S transversally. Moreover, a cut path divides $[0, 1] \times \mathbb{R}$ into two connected components that respectively contain the left and right endpoints of S . For a given cut path γ , we say a point is right of γ when it is in the component containing the right endpoints of S , and we say it is left of γ when it is in the other component. If we extend the local sequences of S to include γ , then each $T(i)$ will include an additional entry, γ , and the local sequence of γ will be a permutation of $[n]$. Let T^γ be the tableau on $[n]$ obtained from T by deleting the prefix that ends in γ from each $T(i)$, and let π^γ be the local sequence of γ .

⁴This is a method from computational geometry introduced by Edelsbrunner and Guibas [6].

Suppose $T^\gamma \neq (\emptyset, \dots, \emptyset)$. We claim that there exists a pair of integers j, k such that $T^\gamma(j) = (k, \dots)$ and $T^\gamma(k) = (j, \dots)$. We call such a pair *adjacent* with respect to γ . Note that j, k must be consecutive in π^γ . To see why such a pair exists, let p be the leftmost point among the crossings of S that are right of γ , and let j and k be the paths crossing at the point p . This means $T^\gamma(j) \neq \emptyset$, and by x -monotonicity, if $T^\gamma(j) = (i, \dots)$ for some $i \neq k$, then paths j and i would cross at a point right of γ and left of p , contradicting our choice of p . Therefore $T^\gamma(j) = (k, \dots)$, and similarly $T^\gamma(k) = (j, \dots)$, so j and k are adjacent with respect to γ .

Suppose $\{j, k\}$ are adjacent with respect to γ with $T^\gamma(j) = (k) \cdot \omega_j$ and $T^\gamma(k) = (j) \cdot \omega_k$. Then, we can perturb γ to obtain a new cut path δ such that $T^\delta(j) = \omega_j$ and $T^\delta(k) = \omega_k$ while $T^\delta(i) = T^\gamma(i)$ for all $i \notin \{j, k\}$. The local sequence of δ is obtained from π by transposing the elements j and k . Call such a modification of the cut path a (j, k) -sweep.

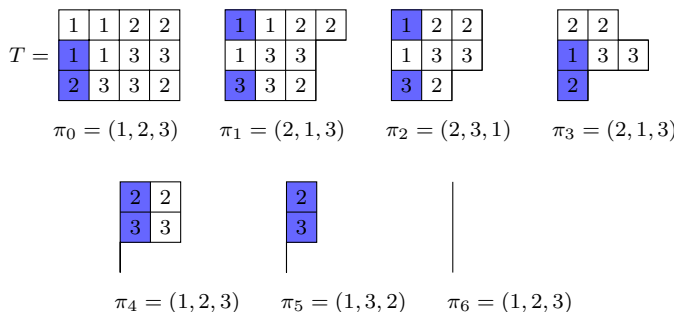


To see why Algorithm 4.2 returns **True** on the geometric tableau T associated to S , start with a cut path γ_0 such that all crossings of S are right of γ_0 . This gives us a tableau $T^{\gamma_0} = T$ and a permutation $\pi^{\gamma_0} = (1, \dots, n)$. The algorithm corresponds to a sequence of sweeps starting from γ_0 in the following way. Let (j, k) be the lexicographically minimal pair that is adjacent with respect to cut path γ_t , and perform a (j, k) -sweep. This produces a new cut path γ_{t+1} with local sequence $\pi^{\gamma_{t+1}}$ and a new tableau $T^{\gamma_{t+1}}$. By the observations above, $T^{\gamma_{t+1}}$ and $\pi^{\gamma_{t+1}}$ are obtained from T^{γ_t} and π^{γ_t} according to a step of Algorithm 4.2, so by induction on t , $\pi = \pi^{\gamma_t}$ and $T = T^{\gamma_t}$ at step t . This procedure can be repeated until all crossings of S are left of a cut path γ_m , which gives $T^{\gamma_m} = (\emptyset, \dots, \emptyset)$. Thus, the algorithm returns **True**.

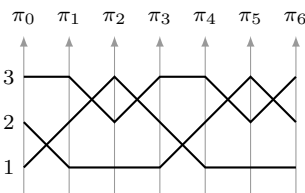
For the converse direction, consider a tableau T on which Algorithm 4.2 returns **True**. We show that T is geometric by constructing a system of paths having T as its associated tableau.⁵ Let (π_0, \dots, π_m) be the sequence of permutations produced by the algorithm. To construct the wiring diagram, start with distinct points on the line $x = 0$ labeled by $[n]$ from bottom to top. Point i will be the left endpoint of path i . At each step of Algorithm 4.2, extend the paths to the right by crossing paths j, k and letting the remaining paths continue horizontally. Notice that at step t a vertical line to the immediate right of this crossing meets the paths in the order of π_t from bottom to top. Furthermore, path i intersects the remaining paths in the order of $T(i)$ from left to right. Hence, T is the tableau associate with S . To make the construction explicit, define path i as the polygonal path with vertices $p_{i,1}, \dots, p_{i,m}$ where $p_{i,t} = m^{-1}(t, h)$ and $i = \pi_t(h)$. \square

⁵This is essentially the same as the “wiring diagram” construction introduced by Goodman [10] as a canonical way of drawing pseudoline arrangements, i.e. 1-crossing systems.

Example 4.4. Below is the state at each step of Algorithm 4.2 on a tableau T . The shaded boxes indicate the adjacent pair at the current step.



Since the final tableau is $(\emptyset, \emptyset, \emptyset)$, it follows from Proposition 4.3 that T is geometric. Plotting the sequence of permutations (π_0, \dots, π_6) produces the following system.



4.2. Valid matchings. We present a similar algorithm which will be conceptually useful. It basically does the same as Algorithm 4.2, except that it disregards the sequence of permutations.

Algorithm 4.5. *The input is a tableau T on $[n]$. The output is True or False.*

```

T' ← T
while T' ≠ (∅, ..., ∅) do
  if There exists a lexicographically minimal pair of integers (j, k) that satisfy
  the following.
    • T'(j) = (k) · ωj.
    • T'(k) = (j) · ωk.
  then
    T'(j) ← ωj.
    T'(k) ← ωk.
  else
    Return False.
  end if
end while
Return True.
    
```

Definition 4.6. Given $T \in [n]^p$, we call a pair of integers $\{j, k\}$ *weakly adjacent* in T if $T(j) = (k, \dots)$ and $T(k) = (j, \dots)$. We say that T has a *valid matching* when the output of Algorithm 4.5 on T is True.

Proposition 4.7. *For $T \in [n]^p$ with $n \geq 3$, T is geometric if and only if T has a valid matching and the restriction T_X is geometric for every 3-subset $X \subset [n]$.*

Proof. One direction is trivial. For the other direction, let T have a valid matching, let T_x be geometric for all 3-subsets $X \subset [n]$, and suppose T is not geometric. By Proposition 4.3, Algorithm 4.2 returns **False**, which means Algorithms 4.2 and 4.5 diverge. Let Y_1, \dots, Y_p be the sequence of weakly adjacent pairs at each step of Algorithm 4.5, let $\pi_0, \pi_1, \dots, \pi_q$ be the sequence of permutations at each step of Algorithm 4.2 prior to divergence, and let $Y_{q+1} = \{j, k\}$. For divergence to occur, $\{j, k\}$ must not be adjacent. That is, there must be some integer i appearing between j and k in π_q . Consider $U = T_{\{i,j,k\}}$.

The subsequence $Y_{t_1}, \dots, Y_{t_m}, Y_{q+1}, \dots$ consisting of all pairs $Y_{t_s} \subset \{i, j, k\}$ defines a valid matching on U , and the order of $\{i, j, k\}$ in the permutation π_{t_s} is determined by sequentially transposing pairs Y_{t_1}, \dots, Y_{t_s} . For a tableau on three elements, there can be at most one weakly adjacent pair of integers. Therefore, the sequence of adjacent integers produced by Algorithm 4.2 on U is exactly Y_{t_1}, \dots, Y_{t_m} , and the sequence of permutations is exactly the order of $\{i, j, k\}$ in $\pi_{t_1}, \dots, \pi_{t_m}$, after which there is no adjacent pair and the algorithm returns **False**. But by Proposition 4.3, Algorithm 4.2 returns **True** on U , since U is geometric. Hence by contradiction, T must be geometric. \square

4.3. Regular tableaux and regular systems. Here we make some observations concerning regular systems and their associated tableaux. Proposition 4.10 concerns the concatenation of regular geometric tableaux and is important in the remaining sections.

Lemma 4.8. *Let S be a system labeled by [3] with signature σ and associated tableau T . Then*

$$T(1) \sim \sigma_{\{x,y\}}, T(2) \sim \sigma_{\{x,z\}}, T(3) \sim \sigma_{\{y,z\}}.$$

Proof. This is just a reformulation of Remark 2.6. \square

Lemma 4.9. *Let S be a k -crossing system and T its associated tableau. Then S is regular if and only if T is regular.*

Proof. Suppose S is labeled by $[n]$. T is regular if and only if for every 3-subset $X \subset [n]$ we have $T_x \sim T_{[3]}$. The claim therefore follows from Lemma 4.8. \square

Given a word α on the alphabet $\{x, y, z\}$, let $X(\alpha)$, $Y(\alpha)$, and $Z(\alpha)$ denote the number of x 's, y 's and z 's in α , respectively.

Proposition 4.10. *Let S be a regular system of size n with signature σ and associated tableaux T_n^ω . If $\sigma = \sigma_1 \cdot \sigma_2$ with $X(\sigma_j) = Y(\sigma_j) = Z(\sigma_j) > 0$, then there exist regular systems S_1 and S_2 of size n with signatures σ_1 and σ_2 . Moreover, if $T_n^{\omega_1}$ and $T_n^{\omega_2}$ are the tableaux associated to S_1 and S_2 , then $T_n^\omega = T_n^{\omega_1 \cdot \omega_2}$.*

Proof. Set $T := T_n^\omega$. By assumption there are positive integers k_1 and k_2 such that $k_j := X(\sigma_j) = Y(\sigma_j) = Z(\sigma_j)$. For every $i \in [n]$, write $T(i) = T_1(i) \cdot T_2(i)$, where $T_1(i)$ consists of the initial $k_1(n - 1)$ entries of $T(i)$, and $T_2(i)$ consists of the final $k_2(n - 1)$ entries of $T(i)$. The sequences $T_1(i)$ and $T_2(i)$ form two tableaux, $T_1 = (T_1(1), \dots, T_1(n))$ and $T_2 = (T_2(1), \dots, T_2(n))$, which satisfy $T = T_1 \cdot T_2$. We first show that T_1 and T_2 are regular.

Let $A = \{i_1, i_2, i_3\} \subset [n]$ be an arbitrary subset with $i_1 < i_2 < i_3$, and consider the restriction T_A . Since T is regular, Lemma 4.8 implies that $T(i_1)_A \sim \sigma_{\{x,y\}}$,

$T(i_2)_A \sim \sigma_{\{x,z\}}$, and $T(i_3)_A \sim \sigma_{\{y,z\}}$. It follows that

$$\begin{aligned} T_1(i_1)_A \cdot T_2(i_1)_A &\sim (\sigma_1)_{\{x,y\}} \cdot (\sigma_2)_{\{x,y\}} \\ T_1(i_2)_A \cdot T_2(i_1)_A &\sim (\sigma_1)_{\{x,z\}} \cdot (\sigma_2)_{\{x,z\}} \\ T_1(i_3)_A \cdot T_2(i_1)_A &\sim (\sigma_1)_{\{y,z\}} \cdot (\sigma_2)_{\{y,z\}}, \end{aligned}$$

and therefore

$$\begin{aligned} T_1(i_1)_A &\sim (\sigma_1)_{\{x,y\}} & , & & T_2(i_1)_A &\sim (\sigma_2)_{\{x,y\}} \\ T_1(i_2)_A &\sim (\sigma_1)_{\{x,z\}} & , & & T_2(i_2)_A &\sim (\sigma_2)_{\{x,z\}} \\ T_1(i_3)_A &\sim (\sigma_1)_{\{y,z\}} & , & & T_2(i_3)_A &\sim (\sigma_2)_{\{y,z\}}. \end{aligned}$$

This proves that T_1 and T_2 are regular, and by Proposition 3.18 there exist words ω_1 and ω_2 in $W_{(ab,cd)}$ such that $T_1 = T_n^{\omega_1}$ and $T_2 = T_n^{\omega_2}$. By Corollary 3.17, $T = T_n^{\omega_1 \cdot \omega_2}$.

We now show that T_1 and T_2 are geometric. Each path of S is involved in $(k_1 + k_2)(n - 1)$ crossings. Cut each path of S at a point strictly between the $k_1(n - 1)$ 'th and the $k_1(n - 1) + 1$ 'st crossings. After suitable homeomorphisms of the plane, we may view this as two distinct systems of paths, S_1 and S_2 , where the crossings of S_1 are precisely the initial $k_1(n - 1)$ crossings of each path of S , and the crossings of S_2 are precisely the final $k_2(n - 1)$ crossings of each path of S . It is easily seen that S_1 and S_2 are regular of size n . Clearly S_1 has signature σ_1 and associated tableau T_1 . Furthermore, if k_1 is even, then S_2 has signature σ_2 and associated tableau T_2 . On the other hand, if k_1 is odd, then the right endpoints of the paths in S_1 appear in reverse order. Thus, if we consider a vertical reflection of S_2 we obtain a regular system of size n which has signature σ_2 and associated tableau T_2 . □

Example 4.11. Consider the following system S of size 4:



By inspection we see that S is regular with signature $\sigma = xyz^3xy^2x$. The associated tableau is T_4^ω where $\omega = (ab)(cd)(cd)(dc)(b^2a^2)$.

1	2	3	3	3	2	2	1	1
1	2	4	4	4	2	2	1	1
1	3	4	3	4	4	3	1	1
2	3	4	2	3	4	4	3	2

Notice that σ can be written as $\sigma = \sigma_1 \cdot \sigma_2$ where $\sigma_1 = xyz$ and $\sigma_2 = z^2xy^2x$. Therefore we get the following regular systems S_1 and S_2 of size 4 with signatures σ_1 and σ_2 . Notice that since $X(\sigma_1) = 1$, S_2 is a vertical reflection of its corresponding part in S .



Their associated tableaux are $T_n^{\omega_1}$ and $T_n^{\omega_2}$ where $\omega_1 = (ab)(cd)$ and $\omega_2 = (cd)(dc)(b^2a^2)$.

1	2	3
1	2	4
1	3	4
2	3	4

3	3	2	2	1	1
4	4	2	2	1	1
3	4	4	3	1	1
2	3	4	4	3	2

5. CHARACTERIZATION OF REGULAR SYSTEMS

We are ready to state our characterization of regular systems (Proposition 5.4). This is given in terms of certain combinatorial conditions on the signatures of k -crossing systems of size 3. We also deduce some consequences of the characterization which will be used in the proofs of Theorems 2.15 and 2.16. The proof of Proposition 5.4 is given at the end of this section.

Definition 5.1. Let σ be the signature of a k -crossing system of size 3. We say that σ is *extendable* if there exists a regular system of size 4 with signature σ .

Remark 5.2. By definition, if S is a regular system with signature σ , then σ is necessarily extendable, since a regular system has size at least 4 and the property of being regular is inherited by all subsystems of size 4. A less obvious fact is that for every extendable signature σ and every integer $n > 4$ there exists a regular system of size n with signature σ . Thus the characterization of a regular system amounts to the characterization of extendable signatures. The proof of this fact will be implicit in our proof of Proposition 5.4.

We start with a few simple reductions. First notice that there are certain invariant symmetries. For instance, a signature σ is extendable if and only if the *reverse* signature, $-\sigma$, is extendable. This is simply the effect of a horizontal (and possibly vertical) reflection of the corresponding system of paths. The other invariant symmetry is that of *interchanging* the x 's and z 's in σ . This corresponds to a vertical reflection of the corresponding system of paths. These symmetries will be used to reduce some of the case analysis in our arguments.

Definition 5.3. A non-empty word $\omega \in \{x, y, z\}^*$ is called *reducible* if there is a non-empty proper initial substring ω' of ω such that $X(\omega') = Y(\omega') = Z(\omega')$. If ω is not reducible, then ω is called *irreducible*.

In view of Proposition 4.10 it suffices to characterize extendable irreducible signatures.

Proposition 5.4. *If σ is an irreducible signature, then σ is extendable if and only if the following conditions hold.*

- (1) *There exists an $\omega \in W_{(ab,cd)}$ such that T_3^ω is the tableau associated to a system with signature σ .*
- (2) *$\sigma = \alpha \cdot \beta \cdot \gamma$ or $\sigma = \gamma \cdot \beta \cdot \alpha$ such that*

$$\begin{pmatrix} X(\alpha) & Y(\alpha) & Z(\alpha) \\ X(\beta) & Y(\beta) & Z(\beta) \\ X(\gamma) & Y(\gamma) & Z(\gamma) \end{pmatrix} = \begin{pmatrix} p & p & 0 \\ q & 0 & p \\ 0 & q & q \end{pmatrix}$$

where p and q are non-negative integers not both equal to 0.

Furthermore, if S is a regular system labeled by $[n]$ with irreducible signature σ and $\omega \in W_{(ab,cd)}$ is the word satisfying condition (1), then T_n^ω is the tableau associated to S .

Remark 5.5. Condition (1) of Proposition 5.4 is a purely combinatorial condition which relates the signature of a regular system to the associated regular tableau. Notice that if there is a word $\omega \in W_{(ab,cd)}$ which satisfies condition (1) of Proposition 5.4, then it is unique and each distinct letter occurs the same number of times. We call this the word *associated* to the irreducible signature σ . A signature and its

associated word are related by Lemma 4.8, and Proposition 3.13 implies that a signature σ has an associated word if and only if the following conditions hold:

- $\sigma_{\{x,y\}}$ is balanced and $\langle \sigma_{\{x,y\}} \rangle$ is a refinement of $\exp_z(\sigma_{\{x,z\}})$.
- $\sigma_{\{y,z\}}$ is balanced and $\langle \sigma_{\{y,z\}} \rangle$ is a refinement of $\exp_x(\sigma_{\{x,z\}})$.

Remark 5.6. Condition (2) of Proposition 5.4 is a geometric condition related to the extendability. As soon as an irreducible signature σ has an associated word $\omega \in W_{(ab,cd)}$, we automatically obtain the tableaux T_n^ω . For $n = 3$ this is the tableau associated to a system with signature σ , and for all $n \geq 4$ these tableau are regular. It turns out that the tableaux T_n^ω are geometric for all $n \geq 4$ if and only if condition (2) holds.

Proposition 5.4 provides a simple way to determine whether an irreducible signature is extendable. Condition (1) tells us that the restrictions $\sigma_{\{x,y\}}$ and $\sigma_{\{y,z\}}$ factor into balanced blocks, and the restriction $\sigma_{\{x,z\}}$ tells us how these balanced blocks should be arranged to form the associated word ω . We illustrate this with several examples.

Example 5.7. Let $\sigma = xy^4x^3z^4x^2zy^2z$. It is easily verified that σ is irreducible and satisfies Proposition 2.8, which implies σ is the signature of a system. Now we show that σ satisfies the conditions of Proposition 5.4.

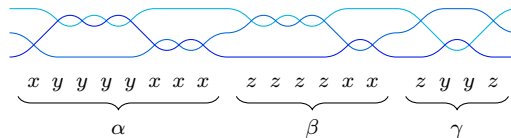
For condition (1), we find the associated word ω by first factoring $\sigma_{\{x,y\}}$ and $\sigma_{\{y,z\}}$ into balanced blocks:

$$\sigma_{\{x,y\}} = (xy)(y^3x^3)(x^2y^2) \sim \omega_{\{c,d\}} \quad \text{and} \quad \sigma_{\{y,z\}} = (y^4z^4)(zy)(yz) \sim \omega_{\{a,b\}}.$$

Then interlace the corresponding blocks of $\omega_{\{c,d\}}$ and $\omega_{\{a,b\}}$ according to $\sigma_{\{x,z\}}$ as in Proposition 3.13:

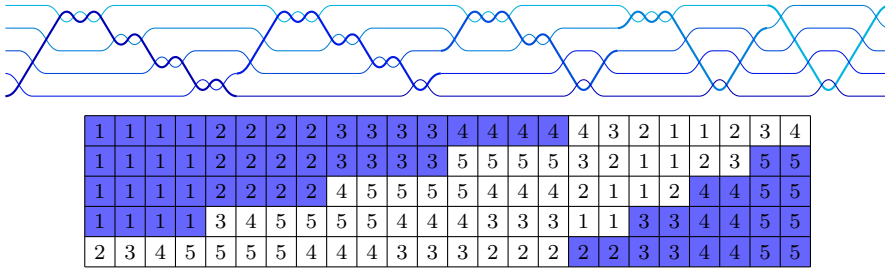
$$\begin{array}{cccccc} \sigma_{\{x,z\}} = & x^4 & z & z^3 & x & x & z^2 \\ & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ \omega = & (a^4b^4) & (cd) & (d^3c^3) & (ba) & (ab) & (c^2d^2). \end{array}$$

For condition (2), σ factors as shown below:



Informally, we may interpret paths in a regular system as “taking turns” crossing the other paths, and interpret the factorization in condition (2) as indicating how each path crosses other paths on its turn. This can be seen more precisely in the proof of Proposition 5.4, but for now just consider the factors in this example. Here, $\sigma = \alpha \cdot \beta \cdot \gamma$ indicates that the paths take turns in order $1, \dots, n$. The term $\alpha = (xy)(y^3x^3)$ indicates that path i crosses paths $[n]_i^+$ on its turn first in ascending order one at a time and then in descending order three at a time. The term $\gamma = (zy)(yz)$ indicates that path i crosses paths $[n]_i^-$ on its turn first in descending and then in ascending order one at a time, and the term $\beta = (z^4x^2)$ indicates that path i first crosses paths $[n]_i^+$, then paths $[n]_i^-$. Below we see a regular system

S' with associated tableau T_5^ω . “Turns” are indicated by thickened paths in the system and by unshaded boxes in the tableau.



Compare path 3 of S' on its turn to the terms α and γ . This example is particularly simple in that there is no interlacing of the order path i crosses paths $[n]_i^+$ and $[n]_i^-$ on its turn, and path i crosses all other paths an even number of times, returning to its original position in order among the other paths at the end of its turn.

Example 5.8. Let $\sigma = xy^2x^2yzx^4z^3y^4z^3$ and observe that σ is the signature of a system of paths and is also irreducible.

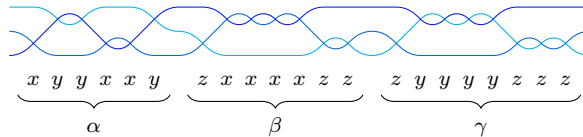
For condition (1), first notice that $\sigma_{\{x,y\}}$ and $\sigma_{\{y,z\}}$ can be written as

$$\sigma_{\{x,y\}} = (xy)(yx)(xy)(x^4y^4) \quad \text{and} \quad \sigma_{\{y,z\}} = (y^3z^3)(zy)(y^3z^3)$$

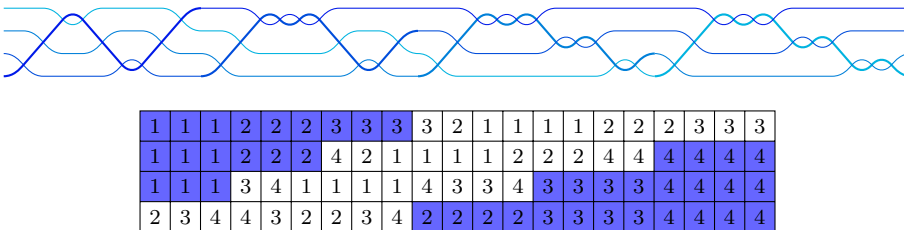
and that $\sigma_{\{x,z\}} = x^3zx^4z^6$. This gives us the associated word

$$\omega = (a^3b^3)(cd)(ba)(a^3b^3)(dc)(cd)(c^4d^4).$$

For condition (2), σ factors as shown below:



Below we see a regular system with associated tableau T_4^ω :



Example 5.9. Here we give an example of a signature which satisfies condition (1) of Proposition 5.4, but not condition (2). Consider the following system:



Its signature is $\sigma = xy^2x^4yz^4y^2z$, and the associated word is $\omega = (a^3b^3)(ba)(ab)(cd)(dc)(c^3d^3)$.

We have $\sigma_{\{x,y\}} \sim \omega_{\{c,d\}}$, $\sigma_{\{x,z\}} \sim \omega_{\{a,c\}}$, and $\sigma_{\{y,z\}} \sim \omega_{\{a,b\}}$, which shows that condition (1) is satisfied.

We now show that σ is not extendable by applying Algorithm 4.5 to the tableau T_4^ω . By Proposition 4.7, T_4^ω is geometric if and only if it has a valid matching. In the figure below, the part of T_4^ω which is shaded consists of the weakly adjacent pairs which get deleted during Algorithm 4.5. Observe that after deleting these entries, the remaining tableau has no weakly adjacent pairs. This proves that ω is not extendable.

1	1	1	2	2	2	3	3	3	3	2	1	1	2	3
1	1	1	2	2	2	2	1	1	2	4	4	4	4	4
1	1	1	1	1	3	4	4	3	3	3	3	4	4	4
2	3	4	4	3	2	2	2	2	3	3	3	4	4	4

In Proposition 5.4, it is assumed that σ is the signature of a system. An arbitrary word $\sigma \in \{x, y, z\}^*$ must satisfy the parity conditions of Proposition 2.8 in order to be a signature, and this imposes additional structure on the words in $W_{(ab,cd)}$ associated to extendable signatures. This structure is described below in Lemma 5.11. It will be crucial for the proof of Proposition 5.4 as well as for the proofs of Theorems 2.15 and 2.16. First, some notions need to be introduced.

Recall the languages $B_{(ab)}$ and $B_{(cd)}$ consisting of balanced blocks on letters $\{a, b\}$ and $\{c, d\}$, respectively (defined in the paragraph preceding Lemma 3.5). Let $\overline{B}_{(ab)}$ and $\overline{B}_{(cd)}$ denote the subsets consisting of *odd* balanced blocks, meaning balanced blocks of the form $(a^k b^k)$, $(b^k a^k)$ and $(c^k d^k)$, $(d^k c^k)$ where k is an odd positive integer. Let $U_{(ab,cd)}$ denote the set of words, $\omega = \omega_1 \cdots \omega_k$, where each $\omega_i \in \overline{B}_{(ab)} \cup \overline{B}_{(cd)}$.

Definition 5.10. A word $\omega \in W_{(ab,cd)}$ is called *well-balanced* if there exist words ω_1 and $\omega_2 \in B_{(ab)} \cup B_{(cd)}$ such that

- $\omega = \omega_1 \cdot \omega' = \omega_2 \cdot \omega''$,
- ω_1 contains an *even* number of a 's and an *odd* number of c 's,
- ω_2 contains an *odd* number of a 's and an *even* number of c 's.

Lemma 5.11. Let σ be an irreducible signature which satisfies conditions (1) and (2) of Proposition 5.4 where σ can be written as $\sigma = \alpha \cdot \beta \cdot \gamma$, and let $\omega \in W_{(ab,cd)}$ be the associated word.

- If $p > 0$ and $q = 0$, then ω is of the form

$$(a^p b^p) \cdot \delta$$

where $\delta = \delta_1 \delta_2 \cdots$, and each $\delta_i \in \overline{B}_{(cd)}$.

- If p and q are both non-zero, then ω is well-balanced and of the form

$$\omega = (a^p b^p) \cdot \delta \cdot (c^q d^q)$$

where $\delta \in U_{(ab,cd)}$.

Remark 5.12. Notice that up to reversal of σ and interchanging x 's and z 's, Lemma 5.11 covers all possible irreducible signatures satisfying the conditions of Proposition 5.4.

Proof of Lemma 5.11. We first consider the case when $p > 0$ and $q = 0$. This means that σ is of the form

$$\sigma = \alpha \cdot z^p,$$

where $X(\alpha) = Y(\alpha) = p$ and $Z(\alpha) = 0$. It follows that $\sigma_{\{y,z\}} = y^p z^p$ and $\sigma_{\{x,z\}} = x^p z^p$. By condition (1) of Proposition 5.4, $\sigma_{\{x,y\}}$ factors into balanced blocks. By Proposition 2.8, there are *odd* positive integers p_i with $\sum p_i = p$ such that

$$\sigma_{\{x,y\}} = (x^{p_1} y^{p_1})(y^{p_2} x^{p_2}) \dots$$

where the blocks of $\sigma_{\{x,y\}}$ appear in an alternating pattern, meaning that we have $(x^{p_i} y^{p_i})$ for odd i , and $(y^{p_i} x^{p_i})$ for even i . Therefore we have

$$\omega = (a^p b^p)(c^{p_1} d^{p_1})(d^{p_2} c^{p_2}) \dots,$$

which is what we wanted to show.

Now suppose both p and q are positive. The assumption that $\sigma = \alpha \cdot \beta \cdot \gamma$ implies that σ starts with the letter x and ends with the letter z . It follows that the initial balanced block of $\sigma_{\{y,z\}}$ equals $(y^p z^p)$ and the last balanced block of $\sigma_{\{x,y\}}$ equals $(x^q y^q)$. This implies that ω is of the form

$$\omega = (a^p b^p) \cdot \delta \cdot (c^q d^q)$$

where $\delta \in W_{(ab,cd)}$. We still need to argue that δ is comprised of odd blocks, that is, $\delta \in U_{(ab,cd)}$. By applying Proposition 2.8, as before, it follows that α can be written as

$$\alpha = (x^{p_1} y^{p_1})(y^{p_2} x^{p_2}) \dots$$

where the p_i are odd positive integers with $\sum p_i = p$ and the blocks appear in an alternating pattern. By symmetry, the same argument shows that γ can be written as

$$\gamma = \dots (z^{q_2} y^{q_2})(y^{q_1} z^{q_1})$$

where the q_i are odd positive integers with $\sum q_i = q$ and the blocks appear in an alternating pattern. This implies that $\sigma_{\{x,y\}}$ and $\sigma_{\{y,z\}}$ can be written as

$$\sigma_{\{x,y\}} = (x^{p_1} y^{p_1})(y^{p_2} x^{p_2}) \dots (x^q y^q) \text{ and } \sigma_{\{y,z\}} = (y^p z^p) \dots (z^{q_2} y^{q_2})(y^{q_1} z^{q_1}).$$

Since $\sigma_{\{x,y\}} \sim \omega_{\{c,d\}}$ and $\sigma_{\{y,z\}} \sim \omega_{\{a,b\}}$, this proves that $\delta \in U_{(ab,cd)}$. It remains to prove that ω is well-balanced.

Consider the case when p is odd. The block $(a^p b^p) \in B_{(ab)}$ is an initial substring of ω which contains an odd number of a 's and an even number of c 's. Next, notice that the last block of α must be $(x^{p_j} y^{p_j})$ for some odd number j . Since p_j is odd, Proposition 2.8 implies that the first occurrence of the letter x in β must be preceded by an *odd* number of z 's. Thus β can be written as

$$\beta = z^m x \dots$$

where m is *odd* and strictly less than p , or else σ is not irreducible. Therefore the restriction $\sigma_{\{x,z\}}$ can be written as

$$\sigma_{\{x,z\}} = x^p z^m x \dots,$$

which implies that there is an initial substring of ω ,

$$(a^p b^p) \cdot \delta_1 \cdot \delta_2 \dots \delta_{2k-1} \cdot \delta_{2k}$$

where $\delta_1, \dots, \delta_{2k-1} \in \overline{B}_{(cd)}$ and $\delta_{2k} \in \overline{B}_{(ab)}$. This initial substring has an even number of a 's and precisely m c 's, and therefore ω is well-balanced.

Now consider the case when p is even. Then the last block of α must be of the form $(y^{p_i} x^{p_i})$. Since p_i is odd, Proposition 2.8 implies that the first occurrence of

the letter z in β must be preceded by an *even* number of x 's. Thus β can be written as

$$\beta = x^m z \dots$$

where m is an *even* non-negative integer. We now distinguish the cases whether $m = 0$ or not.

If m is positive, then the restriction of $\sigma_{\{x,z\}}$ can be written as

$$\sigma_{\{x,z\}} = x^{p+m} z \dots,$$

which implies that there is an initial substring of ω ,

$$(a^p b^p) \cdot \delta_1 \dots \delta_{2k} \cdot \delta_{2k+1}$$

where $\delta_1, \dots, \delta_{2k} \in \overline{B}_{(ab)}$ and $\delta_{2k+1} \in \overline{B}_{(cd)}$. This initial substring contains an even number of a 's and an odd number of c 's. Furthermore, the initial substring

$$(a^p b^p) \cdot \delta_1$$

contains an odd number of a 's and an even number of c 's, and therefore ω is well-balanced.

Finally, suppose that $m = 0$. Then Proposition 2.8 implies that β can be written as

$$\beta = z^k x \dots$$

where k is a positive *even* integer which is strictly less than p , or else σ is not irreducible. Therefore the restriction $\sigma_{\{x,z\}}$ can be written as

$$\sigma_{\{x,z\}} = x^p z^k x \dots,$$

which implies that there is an initial substring of ω ,

$$(a^p b^p) \cdot \delta_1 \dots \delta_{2j} \cdot \delta_{2j+1}$$

where $\delta_1, \dots, \delta_{2j} \in \overline{B}_{(cd)}$, and $\delta_{2j+1} \in \overline{B}_{(ab)}$. This initial substring contains an even number of c 's and an odd number of a 's. Furthermore, the initial substring

$$(a^p b^p) \cdot \delta_1$$

contains an even number of a 's and an odd number of c 's, and therefore ω is well-balanced. □

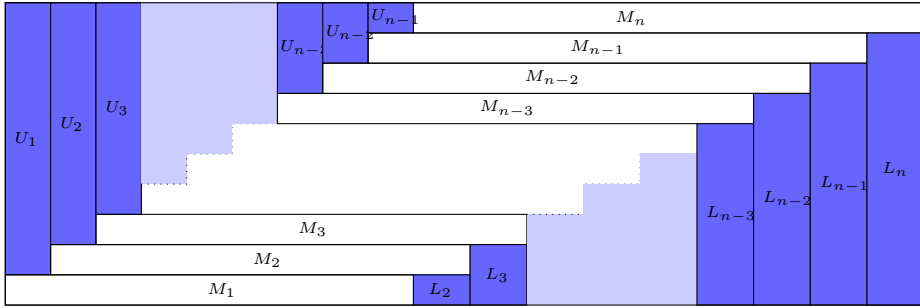
Proof of Proposition 5.4. We first show the sufficiency of the conditions of Proposition 5.4. By Lemma 5.11 we may assume that the associated $\omega \in W_{(ab,cd)}$ can be written as $\omega = (a^p b^p) \cdot \delta \cdot (c^q d^q)$ where $\delta \in U_{(ab,cd)}$. Notice that there are q a 's and p c 's in δ , and that the first case of Lemma 5.11 is obtained by setting $q = 0$. We will show that for any $n \geq 4$, the tableau T_n^ω has a valid matching. By Proposition 4.7 this will imply that T_n^ω is geometric and therefore that σ is extendable.

The tableau T_n^ω comes with a very particular structure. Consider its rows which are given by

$$T_n^\omega(i) = \varphi_{i,n}(\omega) = \varphi_{i,n}(a^p b^p) \cdot \varphi_{i,n}(\delta) \cdot \varphi_{i,n}(c^q d^q).$$

Notice that the term $\varphi_{i,n}(\delta)$ has length $(n-i)p + (i-1)q$ with precisely p entries of each number in $[n]_i^+$ and precisely q entries of each number in $[n]_i^-$. If we view the tableau by its geometric representation we see that we can decompose into parts

as indicated in the figure below:



For $1 \leq i \leq n - 1$, box U_i has height $n - i$ and width p and all its entries equal i . For $1 \leq i \leq n$, box M_i is the part of row i and the entries are given by $\varphi_{i,n}(\delta)$. For $1 < i \leq n$, box L_i has height $i - 1$ and width q and all its entries equal i .

It is now easy to see that T_n^ω has a valid matching. Each entry in M_1 is weakly adjacent to a corresponding entry in U_1 , and therefore Algorithm 4.5 will delete all entries in $U_1 \cup M_1$. Once these are deleted, we see that each entry in M_2 is weakly adjacent to an entry in $U_2 \cup L_2$, and therefore Algorithm 4.5 will delete all entries of $U_2 \cup M_2 \cup L_2$. In general, for $i < n$, each entry in M_i will be weakly adjacent to an entry in $U_i \cup L_i$, and Algorithm 4.5 will delete all entries of $U_i \cup M_i \cup L_i$. Finally, each entry in M_n will be weakly adjacent to an entry in L_n . Notice that in the case when $q = 0$, M_n and every L_i are empty, and each entry in M_i will be weakly adjacent to a member in U_i . This shows that T_n^ω has a valid matching.

We now establish the necessity of condition (1). Suppose σ is an extendable signature. Let S be a regular system of size 4 with signature σ , and let T be the associated tableau. By Lemma 4.9, T is regular, and by Proposition 3.18 there exists an $\omega \in W_{(ab,cd)}$ such that $T = T_4^\omega$. Clearly T_3^ω is the tableau associated to a system with signature σ .

We are ready to show the necessity of condition (2). Suppose σ is irreducible and extendable. Therefore condition (1) holds, and σ satisfies the refinement conditions stated in Remark 5.5. The basic strategy is to factor σ into maximal parts, each consisting of only two of the three letters, and use the refinement conditions to control the distribution of the letters. Throughout we also use the parity conditions from Proposition 2.8. Let

$$\langle \sigma_{\{x,y\}} \rangle = (a_1, a_2, \dots) \quad , \quad \langle \sigma_{\{y,z\}} \rangle = (c_1, c_2, \dots)$$

$$\exp_z(\sigma_{\{x,z\}}) = (m_1, m_2, \dots) \quad , \quad \exp_x(\sigma_{\{x,z\}}) = (n_1, n_2, \dots)$$

and write $\sigma_{\{x,y\}}$ and $\sigma_{\{y,z\}}$ as

$$\sigma_{\{x,y\}} = \alpha_1 \alpha_2 \cdots \quad , \quad \sigma_{\{y,z\}} = \gamma_1 \gamma_2 \cdots$$

where $\alpha_i \in B_{(xy)}$ and $\gamma_i \in B_{(yz)}$. Up to symmetry we may assume that σ starts with an x , and therefore $\alpha_1 = (x^{a_1} y^{a_1})$. Let α be the set of blocks of $\sigma_{\{x,y\}}$ which appear in σ before the first z . We split into cases depending on whether α is empty or not.

Case: $\alpha = \emptyset$. Note that this implies $0 < n_1 \leq a_1$, and we claim that $\sigma = x^q \cdot \sigma_{\{y,z\}}$. If $n_1 < a_1$, then σ can be written as

$$\sigma = x^{n_1} z^{m_1} x \cdots$$

and consequently $n_1 < a_1 \leq m_1 \leq c_1$, which contradicts the assumption that $\langle \sigma_{\{y,z\}} \rangle$ is a refinement of $\exp_x(\sigma_{\{x,z\}})$. Therefore $n_1 = a_1 \leq m_1$, and since $c_1 + c_2 + \dots + c_j = n_1$ for some $j \geq 1$, we conclude that σ can be written as

$$\sigma = x^{n_1} \cdot (\gamma_1 \gamma_2 \cdots \gamma_j).$$

Case: $\alpha \neq \emptyset$. Let $\alpha = \alpha_1 \alpha_2 \cdots \alpha_j$. We note that a_i is odd for all $1 \leq i \leq j$, that $\alpha_i = (x^{a_i} y^{a_i})$ for odd $i \leq j$, and that $\alpha_i = (y^{a_i} x^{a_i})$ for even $i \leq j$. We now split further into subcases depending on whether j is even or odd.

Subcase: j odd. We claim that α_j is directly followed by z . If not, then $\alpha_{j+1} = (y^{a_{j+1}} x^{a_{j+1}})$ and the first z occurs within this block, but this implies $n_1 < c_1$, contradicting the assumption that $\langle \sigma_{\{y,z\}} \rangle$ is a refinement of $\exp_x(\sigma_{\{x,z\}})$. Therefore $c_1 = n_1$ and σ can be written as

$$\sigma = \alpha_1 \alpha_2 \cdots \alpha_j z^c \cdots .$$

If c is even, then z^c is followed by y , implying that $c > n_1$, and therefore σ can be written as

$$\sigma = \alpha_1 \alpha_2 \cdots \alpha_j z^{n_1} z \cdots ,$$

contradicting the assumption that σ is irreducible. Consequently, $c = m_1$ is odd and $m_1 \leq n_1$. If $m_1 = n_1$, then σ can be written as

$$\sigma = \alpha_1 \alpha_2 \cdots \alpha_j z^{n_1}$$

and we are done. Therefore assume that $m_1 < n_1$, which implies $\alpha_{j+1} = (x^{a_{j+1}} y^{a_{j+1}})$. For some $i > 1$ we have $m_1 + m_2 + \dots + m_i = n_1$, where n_1 and m_1 are odd, m_2, \dots, m_i are even. Therefore, $a_{j+1} = n_2 + n_3 + \dots + n_{i+1} \leq m_{i+1}$ and there is a $k \geq i + 1$ for which the sequence (c_2, c_3, \dots, c_k) is a refinement of $(n_2, n_3, \dots, n_{i+1})$. It then follows that σ can be written as

$$\sigma = (\alpha_1 \alpha_2 \cdots \alpha_j)(z^{m_1} x^{n_2} z^{m_2} x^{n_3} \cdots z^{m_i} x^{n_{i+1}})(\gamma_2 \gamma_3 \cdots \gamma_k).$$

Subcase: j even. Similarly to the above, we exclude the case that $\sigma = \alpha_1 \alpha_2 \cdots \alpha_j z^{n_1}$ and we are left with the case that $\alpha_{j+1} = (x^{a_{j+1}} y^{a_{j+1}})$, and we must consider two further subcases depending on whether the initial z is preceded by x or y .

Subsubcase: z preceded by x . This is similar to the case when j is odd. For some $i \geq 1$, we have $m_1 + m_2 + \dots + m_i = n_1$, where each z^{m_s} appears before the initial y in α_{j+1} . This induces a partition of the string of x 's in α_{j+1} giving us the exponent sequence $(n_2, n_3, \dots, n_{i'})$ where $i \leq i' \leq i + 1$ and $a_{j+1} = n_2 + n_3 + \dots + n_{i'} \leq m_{i+1}$. Moreover there is a $k \geq i'$ such that (c_2, c_3, \dots, c_k) is a refinement of $(n_2, n_3, \dots, n_{i'})$, which implies that we can write σ as

$$\sigma = (\alpha_1 \alpha_2 \cdots \alpha_j)(x^d z^{m_1} x^{n_2} z^{m_2} \cdots x^{n_i} z^{m_i} x^e)(\gamma_2 \gamma_3 \cdots \gamma_k)$$

where $d \geq 0$ is even, and $e = 0$ with $i' = i$ or $0 < e = n_{i'}$ with $i' = i + 1$.

Subsubcase: z preceded by y . This case is different from the preceding ones and cannot be dealt with by only using the refinement conditions. For instance, the signature in Example 5.9 belongs to this case. We may assume $\alpha_{j+1} = (x^{a_{j+1}} y^{a_{j+1}})$ with a_{j+1} odd and $n_1 = a_1 + a_2 + \dots + a_{j+1}$. We can write σ as

$$\sigma = (\alpha_1 \alpha_2 \cdots \alpha_j) \cdot (x^{a_{j+1}} y^b z^c y \cdots)$$

where b is odd with $0 < b < a_{j+1}$ and $c > 0$ is even. Set $a = a_1 + \dots + a_j$. Now we have $c_1 = a + b < c < n_1 = a + a_{j+1}$,

$$\begin{aligned} \langle \sigma_{\{x,y\}} \rangle &= (a_1, a_2, \dots) & , & \quad \langle \sigma_{\{y,z\}} \rangle = (c_1, c - c_1, \dots) \\ \exp_z(\sigma_{\{x,z\}}) &= (c_1 + d, m_2, \dots) & , & \quad \exp_x(\sigma_y) = (a + a_{j+1}, n_2, \dots). \end{aligned}$$

From this we can get the following partial structure on ω :

$$\omega = (a^{c_1} b^{c_1})(b^{c-c_1} a^{c-c_1}) \dots (e^{a_1} d^{a_1})(d^{a_2} e^{a_2}) \dots (d^{a_j} e^{a_j})(c^{a_{j+1}} d^{a_{j+1}}) \dots .$$

This partial information is sufficient to show that the tableau T_4^ω does not have a valid matching. Since $T_4^\omega(i) = \varphi_{i,4}(\omega)$, we get

$$\begin{aligned} T_4^\omega(4) &= \underbrace{1 \dots 1}_{c_1} \underbrace{2 \dots 2}_{c_1} \underbrace{3 \dots 3}_{c_1} \dots \\ T_4^\omega(3) &= \underbrace{1 \dots 1}_{c_1} \underbrace{2 \dots 2}_{c_1} \dots \\ T_4^\omega(2) &= \underbrace{1 \dots 1}_{n_1} \underbrace{3 \dots 3}_{a_1} \underbrace{4 \dots 4}_{a_1} \dots \\ T_4^\omega(1) &= \underbrace{2 \dots 2}_{a_1} \underbrace{3 \dots 3}_{a_1} \underbrace{4 \dots 4}_{a_1} \underbrace{4 \dots 4}_{a_2} \underbrace{3 \dots 3}_{a_2} \underbrace{2 \dots 2}_{a_2} \\ &\quad \dots \underbrace{4 \dots 4}_{a_j} \underbrace{3 \dots 3}_{a_j} \underbrace{2 \dots 2}_{a_j} \underbrace{2 \dots 2}_{a_{j+1}} \underbrace{3 \dots 3}_{a_{j+1}} \underbrace{4 \dots 4}_{a_{j+1}} \dots . \end{aligned}$$

We now apply Algorithm 4.5 to T_4^ω and identify the positions of the weakly adjacent pairs that get deleted. Recall that the algorithm always deletes leftmost entries.

- The first $3a$ steps delete $3a$ entries from $T_4^\omega(1)$ together with a entries from $T_4^\omega(i)$ for $i = 1, 2, 3$.
- The next a_{j+1} steps delete a_{j+1} entries from $T_4^\omega(1)$ together with a_{j+1} entries from $T_4^\omega(2)$.
- The next b steps delete b entries from $T_4^\omega(1)$ together with b entries from $T_3^\omega(3)$.
- The next a_1 steps delete a_1 entries from $T_4^\omega(2)$ together with a_1 entries from $T_4^\omega(3)$.

At this point there are no weakly adjacent pairs, so Algorithm 4.5 will output a non-empty tableau, and consequently T_4^ω is not geometric and σ is not extendable. This completes the proof of the necessity of condition (2). □

6. PROOF OF THEOREMS 2.15 AND 2.16

Let σ be an extendable irreducible signature with $X(\sigma) = Y(\sigma) = Z(\sigma) = r = p + q$ where p and q are the numbers from Proposition 5.4. We distinguish the following types:

- σ is called *odd* if r is odd and $p = 0$ or $q = 0$.
- σ is called *even* if r is even and $p = 0$ or $q = 0$.
- σ is called *mixed* if $p > 0$ and $q > 0$.

Example 6.1. Consider the signature $\sigma = xyz$. This is an irreducible extendable signature of the form $\sigma = \alpha \cdot \beta \cdot \gamma$, where $\alpha = xy$, $\beta = z$, and $\gamma = \emptyset$. Thus $p = 1$ and $q = 0$. Note that this is not unique, since it also satisfies $p = 0$ and $q = 1$,

where $\beta = x$ and $\gamma = yz$. In either case, σ is an *odd* signature and the associated word is $\omega = (ab)(cd)$. For $n = 4$ we get the following regular system:



Notice that the system is lower convex and that the upper envelope contains only paths 1 and 4.

Now consider the signature $\sigma = xy^2xz^2$. This is an irreducible extendable signature of the form $\sigma = \alpha \cdot \beta \cdot \gamma$, where $\alpha = (xy)(yx)$, $\beta = z^2$, and $\gamma = \emptyset$. Thus $p = 2$ and $q = 0$, and σ is an *even* signature with associated word $\omega = (a^2b^2)(cd)(dc)$. For $n = 4$ we get the following regular system:



Notice that the system is upper convex and that the lower envelope contains only paths 1 and 2.

Finally consider the signature $\sigma = xy^2x^3z^3y^2z$. This is an irreducible extendable signature of the form $\sigma = \alpha \cdot \beta \cdot \gamma$, where $\alpha = (xy)(yx)$, $\beta = x^2z^2$, and $\gamma = (zy)(yz)$. Thus $p = 2$ and $q = 2$, and σ is a *mixed* signature with associated word $\omega = (a^2b^2)(ba)(ab)(cd)(dc)(c^2d^2)$. For $n = 4$ we get the following regular system:



Notice that this system is upper convex and lower convex.

Lemma 6.2. *Let S be a regular system labeled by $[n]$ with irreducible signature σ .*

- (1) *If σ is odd, then one of the envelopes of S contains only paths 1 and n while the other envelope contains every path of S .*
- (2) *If σ is even, then either S is lower convex and the upper envelope of S contains only paths $n - 1$ and n or S is upper convex and the lower envelope of S contains only paths 1 and 2.*
- (3) *If σ is mixed, then S is upper convex and lower convex.*

Proof. Let $\omega \in W_{(ab,cd)}$ be the word associated to σ and let $T = T_n^\omega$ be the tableau associated to S . The local sequence of path i is given by $T(i) = \varphi_{i,n}(\omega)$. By using Lemma 5.11 we can obtain the structure of the local sequences of S and apply Lemma 2.3 to determine the envelopes of S .

Suppose σ is even or odd. Up to symmetry we may assume σ is of the form

$$\sigma = \alpha \cdot z^p$$

where p is a positive integer. By Lemma 5.11, $\omega \in W_{(ab,cd)}$ has the form

$$\omega = (a^p b^p) \cdot \delta_1 \cdot \delta_2 \cdots$$

where each $\delta_i \in \overline{B}_{(cd)}$. Obviously the lower envelope of S contains path 1 and the upper envelope of S contains path n . We claim that the upper envelope also contains path 1. To see this, consider the local sequence of path 1 which is given by

$$T(1) = \varphi_{1,n}(\omega) = \varphi_{1,n}(\delta_1)\varphi_{1,n}(\delta_2) \cdots$$

by Proposition 5.4. Since $\varphi_{1,n}(\delta_1)$ is an initial string of $T(1)$ which contains each number in $\{2, \dots, n\}$ an odd number of times, Lemma 2.3 implies that the upper envelope of S contains path 1.

For every $1 < i \leq n$, the local sequence of path i is given by

$$T(i) = \varphi_{i,n}(\omega) = \varphi_{i,n}(a^p b^p) \varphi_{i,n}(\delta_1) \varphi_{i,n}(\delta_2) \cdots .$$

If p is odd, then $\varphi_{i,n}(a^p b^p)$ is an initial string of $T(i)$ which contains every number in $[n]_i^-$ an odd number of times, and Lemma 2.3 implies that the lower envelope of S contains path i . Furthermore, for $1 < i < n$, we see from the initial string $\varphi_{i,n}(a^p b^p)$ that any initial string of $T(i)$ which contains each number in $[n]_i^+$ an odd number of times also contains each number in $[n]_i^-$ an odd number of times. Therefore the upper envelope of S does not contain path i , by Lemma 2.3. Thus S is lower convex, while the upper envelope of S contains only paths 1 and n .

If p is even, then $\varphi_{i,n}(a^p b^p) \varphi_{i,n}(\delta_1)$ is an initial string of $T(i)$ which contains each number in $[n]_i^-$ an even number of times and each number in $[n]_i^+$ an odd number of times. By Lemma 2.3, the upper envelope of S contains path i . Furthermore, for $2 < i < n$ we see from the initial string $\varphi_{i,n}(a^p b^p)$ that there is no initial string of $T(i)$ which contains each number in $[n]_i^+$ an odd number of times. Therefore Lemma 2.3 implies that the lower envelope of S does not contain path i . Finally, if $i = 2$, then the initial term of $T(2)$ is 1, and therefore the lower envelope of S contains path 2. Thus S is upper convex, while the lower envelope contains only paths 1 and 2.

Up to symmetry this proves (1) and (2). It remains to prove (3).

By Lemma 5.11 we may assume that ω is of the form

$$\omega = (a^p b^p) \cdot \delta \cdot (c^q d^q)$$

where $\delta \in U_{(ab,cd)}$. Obviously the lower envelope of S contains path 1 and the upper envelope of S contains path n , and we want to show that S is upper and lower convex. For this we use the fact that ω is well-balanced. For $1 \leq i \leq n$ the local sequence of path i is given by $T(i) = \varphi_{i,n}(\omega)$.

Let $\omega_1 \in B_{(ab)} \cup B_{(cd)}$ be an initial string of ω which consists of an even number of a 's and an odd number of c 's. If $1 \leq i < n$, then $\varphi_{i,n}(\omega_1)$ is an initial string of $T(i)$ which contains each number in $[n]_i^-$ an even number of times and each number in $[n]_i^+$ an odd number of times. By Lemma 2.3, the upper envelope of S contains path i . Hence S is upper convex.

Let $\omega_2 \in B_{(ab)} \cup B_{(cd)}$ be an initial string of ω which consists of an odd number of a 's and an even number of c 's. If $1 < i \leq n$, then $\varphi_{i,n}(\omega_2)$ is an initial string of $T(i)$ which contains each number in $[n]_i^-$ an odd number of times and each number in $[n]_i^+$ an even number of times. By Lemma 2.3, the lower envelope of S contains path i . Hence S is lower convex. □

Proof of Theorem 2.15. Let S be a regular system of size n with signature $\sigma = \sigma_1 \cdot \sigma'$ where σ_1 is irreducible. By Proposition 4.10 there is a regular system S_1 of size n with signature σ_1 , and the paths which appear on the envelopes of S_1 also appear on the corresponding envelopes of S . The claim that S is upper or lower convex follows from Lemma 6.2. □

Proof of Theorem 2.16. Suppose S is regular of size n and that S is not upper convex. Let $\sigma = \sigma_1 \cdots \sigma_m$ be the signature of S where each σ_i is irreducible. It

follows from Proposition 4.10 that the upper envelope of S can be determined from the envelopes corresponding to each individual σ_i , which in turn is determined by Lemma 6.2. This immediately implies that none of the σ_i are mixed. Therefore we may assume that σ is the concatenation of odd and even signatures.

Next, observe that if every σ_i is odd or every σ_i is even, then the upper envelope of S contains at most two distinct paths. If every σ_i is odd, then only paths 1 and n are contained in the upper envelope of S (by (1) of Lemma 6.2). On the other hand, if every σ_i is even, then only paths $n - 1$ and n are contained in the upper envelope of S (by (2) of Lemma 6.2). This proves (1) of the theorem, for if S is 2-crossing, then σ cannot contain both an even and an odd signature.

The argument above tells us that the only way we can have more than two but not all paths in the upper envelope of S is when σ is the concatenation of both odd and even signatures. Suppose that this is the case.

Suppose that σ_i is the only even signature. In this case the upper envelope of S will contain three paths. To see this note that (2) of Lemma 6.2 implies that an even signature contributes only the two topmost paths to the upper envelope of S . These may be either paths 1 and 2 or paths $n - 1$ and n depending on the number of odd signatures which precede σ_i in σ . The odd signatures, on the other hand, will always contribute only the two extreme paths, that is, paths 1 and n . Therefore, if only one of the σ_i is even while all the other are odd, then the upper envelope of S contains either paths 1, 2, and n or paths 1, $n - 1$, and n ; see for instance Example 4.11. This proves (2) of the theorem, for if S is 4-crossing and σ is comprised of both odd and even signatures, then σ can contain at most one even signature.

Finally, suppose that there are at least two even signatures, σ_i and σ_j . As before, σ_i and σ_j contribute only the two topmost paths to the upper envelope of S . But if σ_i and σ_j are separated in σ by an odd number of odd signatures, then the order of the paths will switch between them. Therefore one of them will contribute paths 1 and 2, while the other will contribute paths $n - 1$ and n to the upper envelope of S . The odd signatures still only contribute paths 1 and n . Therefore the upper envelope of S may contain paths 1, 2, $n - 1$ and n , but no more. This proves (3) of the theorem. \square

7. REMARKS AND OPEN PROBLEMS

7.1. As we mentioned in the introduction there exist arbitrarily large families of convex bodies such that any two members have six common tangents, any four members are in convex position, but no five members are in convex position. To construct such a family we apply the same technique as was used in [5]. It relies on the fact that if f is a 2π -periodic real C^2 -smooth function such that $f(t) + f''(t) > 0$ holds for all t , then f is the support function of a convex body. (See for instance Lemma 2.2.3 in [12].) Hence, for any 2π -periodic real C^2 -smooth function f , there exists a constant c_0 such that $f + c$ is the support function of a body for all $c > c_0$.

Now consider any system where each pair of paths crosses an *even* number of times. By identifying the endpoints, the paths can be represented by 2π -periodic functions, which furthermore can be approximated by C^2 -smooth functions. Consequently, there is a common constant we can add to each of the (smoothened) functions which makes the system the dual of a family of convex bodies. As in the proof of Theorem 1.1, intersections between support curves are in one-to-one

correspondence with common tangents, and subfamilies in convex position correspond to subsystems which are upper convex. Thus the construction we are looking for can be obtained from any system where each pair of paths crosses six times, each subsystem of size four is upper convex, and no subsystem of size five is upper convex. A regular system with signature $\sigma = xyzxy^2xz^3yxzy^2zx^2$ would be such an example.

7.2. We noted in the introduction that Pach and Tóth constructed an arbitrarily large family of *segments* where every triple is in convex position, but no four members are. When we dualize this family via the support function it is easily seen that we obtain a regular system. As explained in the previous remark, every regular system can be realized by a family of convex bodies. A natural conjecture is: Can every regular $2k$ -crossing system be obtained as the dual of a family of convex n -gons? (Here n should depend only on k , but not on the number of paths in the system.)

7.3. Determining the quantitative behavior for $h_k(n)$ in Theorem 1.1 remains one of the main open questions. Our bound is certainly not close to the truth, and it should be remarked that the condition that any m_k members are in convex position changes the nature of the problem significantly from the original Erdős–Szekeres problem for points when $m_k > 3$: If every four points are in convex position, then the whole set is in convex position. For the case of families of *pairwise disjoint* convex bodies, Bisztriczky and Fejes Tóth [3] have showed that if every five members are in convex position, then the corresponding Erdős–Szekeres function is polynomial in n , and Tóth [19] showed that this function is in fact linear.

7.4. It would be interesting to find sharper bounds for the function $S_k(n)$ given in Theorem 1.4. The case $k = 1$ is dual to the classical *cups-caps* theorem of Erdős–Szekeres [7, 8] and it is known that $S_1(n) = \binom{2n-4}{n-2} + 1$. We do not know of any better lower bound than this for $k > 1$. In fact, it is unclear whether there is actually a dependency on the number of crossings. This leads us to conjecture the following generalization of the *cups-caps* theorem: For every $n \geq 3$ there is a minimal positive integer $S(n)$ such that any system of paths of size $S(n)$ where each pair of paths crosses at least once contains a subsystem of size n which is upper or lower convex.

ACKNOWLEDGMENT

The authors express sincere gratitude to an anonymous referee for the meticulous comments which greatly improved the exposition of their results.

REFERENCES

- [1] I. Bárány and G. Károlyi, *Problems and results around the Erdős-Szekeres convex polygon theorem*, Discrete and computational geometry (Tokyo, 2000), Lecture Notes in Comput. Sci., vol. 2098, Springer, Berlin, 2001, pp. 91–105, DOI 10.1007/3-540-47738-1_7. MR2043640
- [2] T. Bisztriczky and G. Fejes Tóth, *A generalization of the Erdős-Szekeres convex n -gon theorem*, J. Reine Angew. Math. **395** (1989), 167–170, DOI 10.1515/crll.1989.395.167. MR983064 (90b:52005)
- [3] T. Bisztriczky and G. Fejes Tóth, *Convexly independent sets*, Combinatorica **10** (1990), no. 2, 195–202, DOI 10.1007/BF02123010. MR1082649 (92c:52005)

- [4] A. Björner, M. Las Vergnas, B. Sturmfels, N. White, and G. M. Ziegler, *Oriented matroids*, 2nd ed., Encyclopedia of Mathematics and its Applications, vol. 46, Cambridge University Press, Cambridge, 1999. MR1744046 (2000j:52016)
- [5] M. G. Dobbins, A. Holmsen, and A. Hubard, *The Erdős-Szekeres problem for non-crossing convex sets*, *Mathematika* **60** (2014), no. 2, 463–484, DOI 10.1112/S0025579314000072. MR3229499
- [6] H. Edelsbrunner and L. J. Guibas, *Topologically sweeping an arrangement*, 18th Annual ACM Symposium on Theory of Computing (Berkeley, CA, 1986), *J. Comput. System Sci.* **38** (1989), no. 1, 165–194, DOI 10.1016/0022-0000(89)90038-X. MR990055 (91d:68124)
- [7] P. Erdős and G. Szekeres, *A combinatorial problem in geometry*, *Compositio Math.* **2** (1935), 463–470. MR1556929
- [8] P. Erdős and G. Szekeres, *On some extremum problems in elementary geometry*, *Ann. Univ. Sci. Budapest. Eötvös Sect. Math.* **3–4** (1960/1961), 53–62. MR0133735 (24 #A3560)
- [9] J. Fox, J. Pach, B. Sudakov, and A. Suk, *Erdős-Szekeres-type theorems for monotone paths and convex bodies*, *Proc. Lond. Math. Soc.* (3) **105** (2012), no. 5, 953–982, DOI 10.1112/plms/pds018. MR2997043
- [10] J. E. Goodman, *Proof of a conjecture of Burr, Grünbaum, and Sloane*, *Discrete Math.* **32** (1980), no. 1, 27–35, DOI 10.1016/0012-365X(80)90096-5. MR588905 (82b:51005)
- [11] J. E. Goodman, *Pseudoline arrangements*, *Handbook of discrete and computational geometry*, CRC Press Ser. Discrete Math. Appl., CRC, Boca Raton, FL, 1997, pp. 83–109. MR1730161
- [12] H. Groemer, *Geometric applications of Fourier series and spherical harmonics*, *Encyclopedia of Mathematics and its Applications*, vol. 61, Cambridge University Press, Cambridge, 1996. MR1412143 (97j:52001)
- [13] A. Hubard, L. Montejano, E. Mora, and A. Suk, *Order types of convex bodies*, *Order* **28** (2011), no. 1, 121–130, DOI 10.1007/s11083-010-9156-2. MR2774044 (2012c:52005)
- [14] W. Morris and V. Soltan, *The Erdős-Szekeres problem on points in convex position—a survey*, *Bull. Amer. Math. Soc. (N.S.)* **37** (2000), no. 4, 437–458 (electronic), DOI 10.1090/S0273-0979-00-00877-6. MR1779413 (2001e:52030)
- [15] J. Pach and G. Tóth, *A generalization of the Erdős-Szekeres theorem to disjoint convex sets*, Dedicated to the memory of Paul Erdős, *Discrete Comput. Geom.* **19** (1998), no. 3, Special Issue, 437–445, DOI 10.1007/PL00009361. MR1608885 (99b:52038)
- [16] J. Pach and G. Tóth, *Erdős-Szekeres-type theorems for segments and noncrossing convex sets*, *Geom. Dedicata* **81** (2000), no. 1-3, 1–12, DOI 10.1023/A:1005210202899. MR1772191 (2001g:52027)
- [17] G. Szekeres and L. Peters, *Computer solution to the 17-point Erdős-Szekeres problem*, *ANZIAM J.* **48** (2006), no. 2, 151–164, DOI 10.1017/S144618110000300X. MR2291511 (2008a:52022)
- [18] G. Tóth and P. Valtr, *The Erdős-Szekeres theorem: upper bounds and related results*, *Combinatorial and computational geometry*, *Math. Sci. Res. Inst. Publ.*, vol. 52, Cambridge Univ. Press, Cambridge, 2005, pp. 557–568. MR2178339 (2006k:52044)
- [19] G. Tóth, *Finding convex sets in convex position*, *Combinatorica* **20** (2000), no. 4, 589–596, DOI 10.1007/s004930070010. MR1804829 (2002c:52021)

DEPARTMENT OF MATHEMATICAL SCIENCES, BINGHAMTON UNIVERSITY SUNY, BINGHAMTON,
NEW YORK 13902 – AND – GAIA, POSTECH, POHANG, SOUTH KOREA

E-mail address: `dobbins@postech.ac.kr`

DEPARTMENT OF MATHEMATICAL SCIENCES, KAIST, DAEJEON, SOUTH KOREA

E-mail address: `andreash@kaist.edu`

LABORATOIRE DE L'INSTITUT GASPARD MONGE, UNIVERSITÉ PARIS-EST MARNE-LA-VALLÉE,
PARIS, FRANCE – AND – GEOMETRICA, INRIA SOPHIA-ANTIPOLIS, FRANCE

E-mail address: `alfredo.hubard@inria.fr`