Metadata of the chapter that will be visualized in SpringerLink

Book Title	LATIN 2022: Theoretical Informatics
Series Title	
Chapter Title	Patterns in Ordered (random) Matchings
Copyright Year	2022
Copyright HolderName	Springer Nature Switzerland AG

Corresponding Author	Family Name	**Dudek**
	Particle	
	Given Name	**Andrzej**
	Prefix	
	Suffix	
	Role	
	Division	Department of Mathematics
	Organization	Western Michigan University
	Address	Kalamazoo, MI, USA
	Email	andrzej.dudek@wmich.edu
Author	Family Name	**Grytczuk**
	Particle	
	Given Name	**Jarosław**
	Prefix	
	Suffix	
	Role	
	Division	Faculty of Mathematics and Information Science
	Organization	Warsaw University of Technology
	Address	Warsaw, Poland
	Email	j.grytczuk@mini.pw.edu.pl
Author	Family Name	**Ruciński**
	Particle	
	Given Name	**Andrzej**
	Prefix	
	Suffix	
	Role	
	Division	Department of Discrete Mathematics
	Organization	Adam Mickiewicz University
	Address	Poznań, Poland
	Email	rucinski@amu.edu.pl

Abstract

An *ordered matching* is an ordered graph which consists of vertex-disjoint edges (and have no isolated vertices). In this paper we focus on unavoidable patterns in such matchings. First, we investigate the size of canonical substructures in ordered matchings and generalize the Erdős-Szekeres theorem about monotone sequences. We also estimate the size of canonical substructures in a *random* ordered matching. Then we study *twins*, that is, pairs of order-isomorphic, disjoint sub-matchings. Among other results, we show that every ordered matching of size n contains twins of length $\Omega(n^{3/5})$, but the length of the longest twins in *almost* every ordered matching is $O(n^{2/3})$.

Keywords (separated by '-')	Ordered matchings - Unavoidable patterns - Twins

Patterns in Ordered (random) Matchings

Andrzej Dudek[1](✉), Jarosław Grytczuk[2], and Andrzej Ruciński[3]

[1] Department of Mathematics, Western Michigan University, Kalamazoo, MI, USA
andrzej.dudek@wmich.edu

[2] Faculty of Mathematics and Information Science, Warsaw University of Technology, Warsaw, Poland
j.grytczuk@mini.pw.edu.pl

[3] Department of Discrete Mathematics, Adam Mickiewicz University, Poznań, Poland
rucinski@amu.edu.pl

Abstract. An *ordered matching* is an ordered graph which consists of vertex-disjoint edges (and have no isolated vertices). In this paper we focus on unavoidable patterns in such matchings. First, we investigate the size of canonical substructures in ordered matchings and generalize the Erdős-Szekeres theorem about monotone sequences. We also estimate the size of canonical substructures in a *random* ordered matching. Then we study *twins*, that is, pairs of order-isomorphic, disjoint sub-matchings. Among other results, we show that every ordered matching of size n contains twins of length $\Omega(n^{3/5})$, but the length of the longest twins in *almost* every ordered matching is $O(n^{2/3})$. AQ1

Keywords: Ordered matchings · Unavoidable patterns · Twins

1 Introduction

A graph G is said to be *ordered* if its vertex set is linearly ordered. Let G and H be two ordered graphs with $V(G) = \{v_1 < \cdots < v_m\}$ and $V(H) = \{w_1 < \cdots < w_m\}$ for some integer $m \geq 1$. We say that G and H are *order-isomorphic* if for all $1 \leq i < j \leq m$, $v_iv_j \in G$ if and only if $w_iw_j \in H$. Note that every pair of order-isomorphic graphs is isomorphic, but not vice-versa. Also, if G and H are distinct graphs on the same linearly ordered vertex set V, then G and H are never order-isomorphic, and so all $2^{\binom{|V|}{2}}$ labeled graphs on V are pairwise non-order-isomorphic. It shows that the notion of order-isomorphism makes sense only for pairs of graphs on distinct vertex sets.

One context in which order-isomorphism makes quite a difference is that of subgraph containment. If G is an ordered graph, then any subgraph G' of G can be also treated as an ordered graph with the ordering of $V(G')$ inherited

The first author was supported in part by Simons Foundation Grant #522400. The second author was supported in part by Narodowe Centrum Nauki, grant 2020/37/B/ST1/03298. The third author was supported in part by Narodowe Centrum Nauki, grant 2018/29/B/ST1/00426.

A. Castañeda and F. Rodríguez-Henríquez (Eds.): LATIN 2022, LNCS 13568, pp. 1–13, 2022.
https://doi.org/10.1007/978-3-031-20624-5_33

from the ordering of $V(G)$. Given two ordered graphs, (a "large" one) G and (a "small" one) H, we say that a subgraph $G' \subset G$ is *an ordered copy of H in G* if G' and H are order-isomorphic.

All kinds of questions concerning subgraphs in ordinary graphs can be posed for ordered graphs as well (see, e.g., [11]). For example, in [3], the authors studied Turán and Ramsey type problems for ordered graphs. In particular, they showed that there exists an ordered matching on n vertices for which the (ordered) Ramsey number is super-polynomial in n, a sharp contrast with the linearity of the Ramsey number for ordinary (i.e. unordered) matchings. This shows that it makes sense to study even such seemingly simple structures as ordered matchings. In fact, Jelínek [7] counted the number of matchings avoiding (i.e. not containing) a given small ordered matching.

In this paper we focus exclusively on *ordered matchings*, that is, ordered graphs which consist of vertex-disjoint edges (and have no isolated vertices). For example, in Fig. 1, we depict two ordered matchings, $M = \{\{1,3\},\{2,4\},\{5,6\}\}$ and $N = \{\{1,5\},\{2,3\},\{4,6\}\}$ on vertex set $[6] = \{1,2,\ldots,6\}$ with the natural linear ordering.

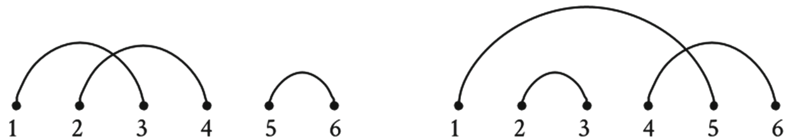

Fig. 1. Exemplary matchings, M and N, of size 3.

A convenient representation of ordered matchings can be obtained in terms of *double occurrence words* over an n-letter alphabet, in which every letter occurs exactly twice as the label of the ends of the corresponding edge in the matching. For instance, our two exemplary matchings can be written as $M = ABABCC$ and $N = ABBCAC$ (see Fig. 2).

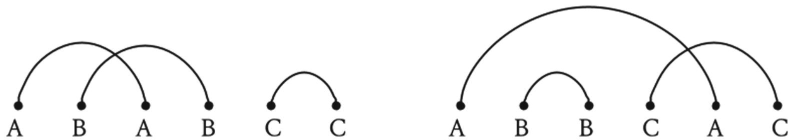

Fig. 2. Exemplary matchings M and N.

Unlike in [7], we study what sub-structures are *unavoidable* in ordered matchings. A frequent theme in both fields, the theory of ordered graphs as well as enumerative combinatorics, are unavoidable sub-structures, that is, *patterns* that appear in every member of a prescribed family of structures. A good example providing everlasting inspiration is the famous theorem of Erdős and Szekeres [5] on monotone subsequences. It states that any sequence $x_1, x_2, \ldots, x_n$ of distinct real numbers contains an increasing or decreasing subsequence of length at least $\sqrt{n}$.

And, indeed, our first goal is to prove its analog for ordered matchings. The reason why the original Erdős-Szekeres Theorem lists only two types of subsequences is, obviously, that for any two elements x_i and x_j with $i < j$ there are just two possible relations: $x_i < x_j$ or $x_i > x_j$. For matchings, however, for every two edges $\{x, y\}$ and $\{u, w\}$ with $x < y$, $u < w$, and $x < u$, there are three possibilities: $y < u$, $w < y$, or $u < y < w$ (see Fig. 3). In other words, every two edges form either *an alignment*, *a nesting*, or *a crossing* (the first term introduced by Kasraoui and Zeng in [8], the last two terms coined in by Stanley [10]). These three possibilities give rise, respectively, to three "unavoidable" ordered *canonical* sub-matchings (*lines*, *stacks*, and *waves*) which play an analogous role to the monotone subsequences in the classical Erdős-Szekeres Theorem.

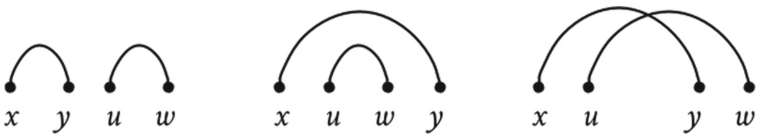

Fig. 3. An alignment, a nesting, and a crossing of a pair of edges.

Informally, lines, stacks, and waves are defined by demanding that every pair of edges in a sub-matching forms, respectively, an alignment, a nesting, or a crossing (see Fig. 5). Here we generalize the Erdős-Szekeres Theorem as follows.

Theorem 1. *Let ℓ, s, w be arbitrary positive integers and let $n = \ell s w + 1$. Then, every ordered matching M on $2n$ vertices contains a line of size $\ell + 1$, or a stack of size $s + 1$, or a wave of size $w + 1$.*

It is not hard to see that the above result is optimal. For example, consider the case $\ell = 5$, $s = 3$, $w = 4$. Take 3 copies of the wave of size $w = 4$: $ABCDABCD$, $PQRSPQRS$, $XYZTXYZT$. Arrange them into a stack-like structure (see Fig. 4):

$$ABCDPQRSXYZTXYZTPQRSABCD.$$

Now, concatenate $\ell = 5$ copies of this structure. Clearly, we obtain a matching of size $\ell s w = 5 \cdot 3 \cdot 4$ with no line of size 6, no stack of size 4, and no wave of size 5.

Also observe that the symmetric case of Theorem 1 implies that M always contains a canonical structure of size at least $n^{1/3}$.

Finally, notice that forbidding an alignment yields to a so called *permutational matching* (for definition see the paragraph after Theorem 4). Permutational matchings are in a one-to-one correspondence with permutations of order n. Moreover, under this bijection waves and stacks in a permutational matching M become, respectively, increasing and decreasing subsequences of

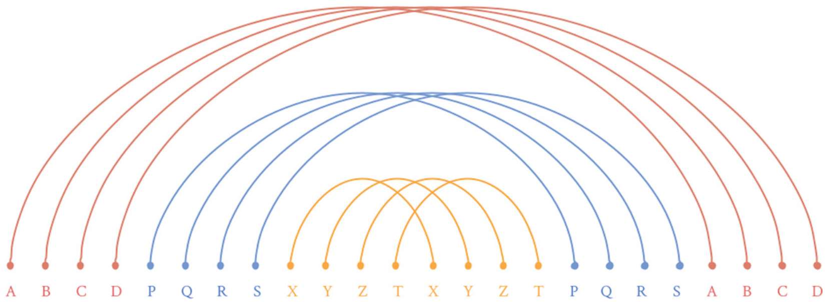

Fig. 4. A stack of waves.

the permutation which is the image of M. Thus, we recover the original Erdős-Szekeres Theorem as a special case of Theorem 1.

We also examine the question of unavoidable sub-matchings for *random* matchings. A random (ordered) matching $\mathbb{RM}_n$ is selected uniformly at random from all $(2n)!/(n!2^n)$ matchings on vertex set $[2n]$. It follows from a result of Stanley (Thoerem 17 in [10]) that *a.a.s.*[1] the size of the largest stack and wave in $\mathbb{RM}_n$ is $(1+o(1))\sqrt{2n}$. In Sect. 2 we complement his result and prove that the maximum size of lines is also about $\sqrt{n}$.

Theorem 2.

(i) A.a.s. the random matching $\mathbb{RM}_n$ *contains no lines of size at least* $e\sqrt{n}$.
(ii) A.a.s. the random matching $\mathbb{RM}_n$ *contains lines of size at least* $\sqrt{n}/8$.

Our second goal is to estimate the size of the largest (ordered) twins in ordered matchings. The problem of twins has been widely studied for other combinatorial structures, including words, permutations, and graphs (see, e.g., [1,9]). We say that two edge-disjoint (ordered) subgraphs G_1 and G_2 of an (ordered) graph G form *(ordered) twins in* G if they are (order-)isomorphic. The size of the (ordered) twins is defined as $|E(G_1)| = |E(G_2)|$. For ordinary matchings, the notion of twins becomes trivial, as every matching of size n contains twins of size $\lfloor n/2 \rfloor$ – just split the matching into two as equal as possible parts. But for ordered matchings the problem becomes interesting. The above mentioned generalization of Erdős-Szekeres Theorem immediately (again by splitting into two equal parts) yields ordered twins of length $\lfloor n^{1/3}/2 \rfloor$. In Sect. 3 we provide much better estimates on the size of largest twins in ordered matchings which, not so surprisingly, are of the same order of magnitude as those for twins in permutations (see [2] and [4]).

2 Unavoidable Sub-matchings

Let us start with formal definitions. Let M be an ordered matching on the vertex set $[2n]$, with edges denoted as $e_i = \{a_i, b_i\}$ so that $a_i < b_i$, for all $i = 1, 2, \ldots, n$,

[1] *Asymptotically almost surely.*

and $a_1 < \cdots < a_n$. We say that an edge e_i is *to the left of* e_j and write $e_i < e_j$ if $a_i < a_j$. That is, in ordering the edges of a matching we ignore the positions of the right endpoints.

We now define the three canonical types of ordered matchings:

- *Line*: $a_1 < b_1 < a_2 < b_2 < \cdots < a_n < b_n$,
- *Stack*: $a_1 < a_2 < \cdots < a_n < b_n < b_{n-1} < \cdots < b_1$,
- *Wave*: $a_1 < a_2 < \cdots < a_n < b_1 < b_2 < \cdots < b_n$.

Assigning letter A_i to edge $\{a_i, b_i\}$, their corresponding double occurrence words look as follows:

- Line: $A_1A_1A_2A_2 \cdots A_nA_n$,
- Stack: $A_1A_2 \cdots A_nA_nA_{n-1} \cdots A_1$,
- Wave: $A_1A_2 \cdots A_nA_1A_2 \cdots A_n$.

Each of these three types of ordered matchings can be equivalently characterized as follows. Let us consider all possible ordered matchings with just two edges. In the double occurrence word notation these are $AABB$ (an alignment), $ABBA$ (a nesting), and $ABAB$ (a crossing). Now a line, a stack, and a wave is an ordered matching in which *every* pair of edges forms, respectively, an alignment, a nesting, and a crossing (see Fig. 5).

Fig. 5. A line, a stack, and a wave of size three.

Consider a sub-matching M' of M and an edge $e \in M \setminus M'$, whose left endpoint is to the left of the left-most edge f of M'. Note that if M' is a line and e and f form an alignment, then $M' \cup \{e\}$ is a line too. Similarly, if M' is a stack and $\{e, f\}$ form a nesting, then $M' \cup \{e\}$ is a stack too. However, an analogous statement fails to be true for waves, as e, though crossing f, may not necessarily cross all other edges of the wave M'. Due to this observation, in the proof of our first result we will need another type of ordered matchings combining lines and waves. We call a matching $M = \{\{a_i, b_i\} : i = 1, \ldots, n\}$ with $a_i < b_i$, for all $i = 1, 2, \ldots, n$, and $a_1 < \cdots < a_n$, *a landscape* if $b_1 < b_2 < \cdots < b_n$, that is, the right-ends of the edges of M are also linearly ordered (a first-come-first-serve pattern). Notice that there are no non-trivial stacks in a landscape. In the double occurrence word notation, a landscape is just a word obtained by a *shuffle* of the two copies of the word $A_1A_2 \cdots A_n$. Examples of landscapes for $n = 4$ are, among others, $ABCDABCD$, $AABCBCDD$, $ABCABDCD$ (see Fig. 6). Now we are ready to prove Theorem 1.

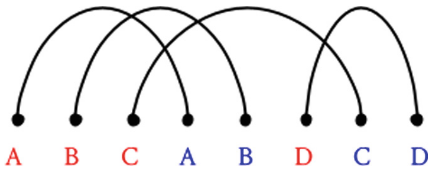

Fig. 6. A landscape of size four.

Proof of Theorem 1. Let M be any ordered matching with edges $\{a_i, b_i\}$, $i = 1, 2, \ldots, n$. Let s_i denote the size of a largest stack whose left-most edge is $\{a_i, b_i\}$. Similarly, let L_i be the largest size of a landscape whose left-most edge is $\{a_i, b_i\}$. Consider the sequence of pairs (s_i, L_i), $i = 1, 2, \ldots, n$. We argue that no two pairs of this sequence may be equal. Indeed, let $i < j$ and consider the two edges $\{a_i, b_i\}$ and $\{a_j, b_j\}$. These two edges may form a nesting, an alignment, or a crossing. In the first case we get $s_i > s_j$, since the edge $\{a_i, b_i\}$ enlarges the largest stack starting at $\{a_j, b_j\}$. In the two other cases, we have $L_i > L_j$ by the same argument. Since the number of pairs (s_i, L_i) is $n > s \cdot \ell w$, it follows that either $s_i > s$ for some i, or $L_j > \ell w$ for some j. In the first case we are done, as there is a stack of size $s + 1$ in M.

In the second case, assume that L is a landscape in M of size at least $p = \ell w + 1$. Let us order the edges of L as $e_1 < e_2 < \cdots < e_p$, accordingly to the linear order of their left ends. Decompose L into edge-disjoint waves, $W_1, W_2, \ldots, W_k$, in the following way. For the first wave W_1, pick e_1 and all edges whose left ends are between the two ends of e_1, say, $W_1 = \{e_1 < e_2 < \ldots < e_{i_1}\}$, for some $i_1 \geqslant 1$. Clearly, W_1 is a true wave since there are no nesting pairs in L. Notice also that the edges e_1 and e_{i_1+1} are non-crossing since otherwise the latter edge would be included in W_1. Now, we may remove the wave W_1 from L and repeat this step for $L - W_1$ to get the next wave $W_2 = \{e_{i_1+1} < e_{i_1+2} < \ldots < e_{i_2}\}$, for some $i_2 \geqslant i_1 + 1$. And so on, until exhausting all edges of L, while forming the last wave $W_k = \{e_{i_{k-1}+1} < e_{i_{k-1}+2} < \ldots < e_{i_k}\}$, with $i_k \geqslant i_{k-1} + 1$. Clearly, the sequence $e_1 < e_{i_1+1} < \ldots < e_{i_{k-1}+1}$ of the leftmost edges of the waves W_i must form a line. So, if $k \geqslant \ell + 1$, we are done. Otherwise, we have $k \leqslant \ell$, and because $p = \ell w + 1$, some wave W_i must have at least $w + 1$ edges. This completes the proof. $\square$

It is not hard to see that the above result is optimal.

Now we change gears a little bit and investigate the size of unavoidable structures in *random* ordered matchings. Let $\mathbb{RM}_n$ be a random (ordered) matching of size n, that is, a matching picked uniformly at random out of the set of all

$$\alpha_n := \frac{(2n)!}{n!2^n}$$

matchings on the set $[2n]$.

Stanley determined very precisely the maximum size of two of our three canonical patterns, stacks, and waves, contained in a random ordered matching.

Theorem 3 (Theorem 17 in *[10]***).** *The largest stack and the largest wave contained in* $\mathbb{RM}_n$ *are each a.a.s. of size* $(1+o(1))\sqrt{2n}$*.*

Our Theorem 2 complements this result by estimating the maximum size of lines. In the proof of Part (ii) of Theorem 2 we will make use of the following lemma that can be easily checked by a standard application of the second moment method, and, therefore, its proof is omitted here. Define the *length* of an edge $\{i,j\}$ in a matching on $[2n]$ as $|j-i|$.

Lemma 1. *Let a sequence* $k=k(n)$ *be such that* $1 \ll k \ll n$*. Then, a.a.s. the number of edges of length at most* k *in* $\mathbb{RM}_n$ *is* $k(1+o(1))$*.*

Proof of Theorem 2. Part (i) is an easy application of the first moment method. Let X_k be a random variable counting the number of ordered copies of lines of size k in $\mathbb{RM}_n$. Then,

$$\mathbb{E}X_k = \binom{2n}{2k}\cdot 1\cdot \frac{\alpha_{n-k}}{\alpha_n} = \frac{2^k}{(2k)!}\cdot\frac{n!}{(n-k)!} \le \frac{2^k}{(2k)!}\cdot n^k \le \frac{2^k}{(2k/e)^{2k}}\cdot n^k = \left(\frac{e^2 n}{2k^2}\right)^k.$$

Thus,

$$\begin{aligned}\Pr(\exists k \ge e\sqrt{n}:\ X_k > 0) &\le \sum_{e\sqrt{n}\le k\le n} \mathbb{E}X_k \\ &\le \sum_{e\sqrt{n}\le k\le n}\left(\frac{e^2 n}{2k^2}\right)^k \le n2^{-e\sqrt{n}} = o(1).\end{aligned}$$

It remains to prove Part (ii). Let $k=\sqrt{n}/2$. Due to Lemma 1, a.a.s. the number of edges of length at most k in $\mathbb{RM}_n$ is at least $\sqrt{n}/4$. We will show that among the edges of length at most k, there are a.a.s. at most $\sqrt{n}/8$ crossings or nestings. After removing one edge from each crossing and nesting we obtain a line of size at least $\sqrt{n}/4-\sqrt{n}/8=\sqrt{n}/8$.

For a 4-element subset $S=\{u_1,u_2,v_1,v_2\}\subset[2n]$ with $u_1<v_1<u_2<v_2$, let X_S be an indicator random variable equal to 1 if both $\{u_1,u_2\}\in\mathbb{RM}_n$ and $\{v_1,v_2\}\in\mathbb{RM}_n$, that is, if S spans a crossing in $\mathbb{RM}_n$. Clearly,

$$\Pr(X_S=1)=\frac{1}{(2n-1)(2n-3)}.$$

Let $X=\sum X_S$, where the summation is taken over all sets S as above and such that $u_2-u_1\le k$ and $v_2-v_1\le k$. Note that this implies that $v_1-u_1\le k-1$. Let $f(n,k)$ denote the number of terms in this sum. We have

$$f(n,k)\le\left(2n(k-1)-\binom{k}{2}\right)\binom{k}{2}\le\left(nk-\frac{1}{2}\binom{k}{2}\right)k^2,$$

as we have at most $2n(k-1)-\binom{k}{2}$ choices for u_1 and v_1 and, once u_1, v_1 have been selected, at most $\binom{k}{2}$ choices of u_2, and v_2. It is easy to see that $f(n,k)=\Omega(nk^3)$. Hence, $\mathbb{E}X=\Omega(k^3/n)\to\infty$, while

$$\mathbb{E}X=\sum_S \mathbb{E}X_S=\frac{f(n,k)}{(2n-1)(2n-3)}\le k^3/4n=\frac{1}{32}\sqrt{n}.$$

To apply Chebyshev's inequality, we need to estimate $\mathbb{E}(X(X-1))$, which can be written as

$$\mathbb{E}(X(X-1))=\sum_{S,S'}\Pr(\{\{u_1,u_2\},\{v_1,v_2\},\{u_1',u_2'\},\{v_1',v_2'\}\}\subset \mathbb{RM}_n),$$

where the summation is taken over all (ordered) pairs of sets $S=\{u_1,u_2,v_1,v_2\}\subset[2n]$ with $u_1<v_1<u_2<v_2$ and $S'=\{u_1',u_2',v_1',v_2'\}\subset[2n]$ with $u_1'<v_1'<u_2'<v_2'$ such that $u_2-u_1\le k$, $v_2-v_1\le k$, $u_2'-u_1'\le k$, and $v_2'-v_1'\le k$. We split the above sum into two sub-sums Σ_1 and Σ_2 according to whether $S\cap S'=\emptyset$ or $|S\cap S'|=2$ (for all other options the above probability is zero). In the former case,

$$\Sigma_1\le\frac{f(n,k)^2}{(2n-1)(2n-3)(2n-5)(2n-7)}=(\mathbb{E}X)^2(1+O(1/n)).$$

In the latter case, the number of such pairs (S,S') is at most $f(n,k)\cdot 4k^2$, as given S, there are four ways to select the common pair and at most k^2 ways to select the remaining two vertices of S'. Thus,

$$\Sigma_2\le\frac{f(n,k)\cdot 4k^2}{(2n-1)(2n-3)(2n-5)}=O(k^5/n^2)=O(\sqrt{n})$$

and, altogether,

$$\mathbb{E}(X(X-1))\le(\mathbb{E}X)^2(1+O(1/n))+O(\sqrt{n})=(\mathbb{E}X)^2+O(\sqrt{n}).$$

By Chebyshev's inequality,

$$\begin{aligned}\Pr(|X-\mathbb{E}X|\ge\mathbb{E}X)&\le\frac{\mathbb{E}(X(X-1))+\mathbb{E}X-(\mathbb{E}X)^2}{(\mathbb{E}X)^2}\\&\le 1+O(1/\sqrt{n})+\frac{1}{\mathbb{E}X}-1=O\left(\frac{1}{\sqrt{n}}\right)\to 0.\end{aligned}$$

Thus, a.a.s. $X\le 2\mathbb{E}X\le\sqrt{n}/16$.

We deal with nestings in a similar way. For a 4-element subset $S=\{u_1,u_2,v_1,v_2\}\subset[2n]$ with $u_1<v_1<v_2<u_2$, let Y_S be an indicator random variable equal to 1 if both $\{u_1,u_2\}\in\mathbb{RM}_n$ and $\{v_1,v_2\}\in\mathbb{RM}_n$, that is, if S spans a nesting in $\mathbb{RM}_n$. Further, let $Y=\sum Y_S$, where the summation is taken over all sets S as above and such that $u_2-u_1\le k$ and $v_2-v_1\le k$. It is

crucial to observe that, again, $\mathbb{E}Y \leq k^3/n = \sqrt{n}/32$. Indeed, this time there are at most $2nk - \binom{k+1}{2}$ choices for u_1 and u_1 and, once u_1, u_1 have been selected, at most $\binom{k}{2}$ choices of v_1, and v_2, while the probability of both pairs appearing in $\mathbb{RM}_n$ remains the same as before. The remainder of the proof goes mutatis mutandis.

We conclude that a.a.s. the number of crossings and nestings of length at most k in $\mathbb{RM}_n$ is at most $\sqrt{n}/8$ as was required. $\square$

3 Twins

Recall that by *twins* in an ordered matching M we mean any pair of disjoint, order-isomorphic sub-matchings M_1 and M_2 and that their *size* is defined as the number of edges in just one of them. For instance, the matching $M = AABCDDEBCFGHIHEGFI$ contains twins $M_1 = BCDDBC$ and $M_2 = EFHHEF$ of size three (see Fig. 7).

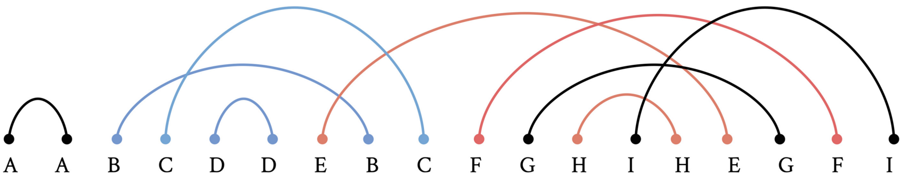

Fig. 7. Twins of size 3 with pattern $XYZZXY$.

Let $t(M)$ denote the maximum size of twins in a matching M and $t_{\text{match}}(n)$ – the minimum of $t(M)$ over all matchings on $[2n]$.

We first point to a direct correspondence between twins in permutations and ordered twins in a certain kind of matchings. By a *permutation* we mean any finite sequence of pairwise distinct positive integers. We say that two permutations $(x_1, \ldots, x_k)$ and $(y_1, \ldots, y_k)$ are *similar* if their entries preserve the same relative order, that is, $x_i < x_j$ if and only if $y_i < y_j$ for all $1 \leqslant i < j \leqslant k$. Any two similar and disjoint sub-permutations of a permutation π are called *twins*. For example, in permutation

$$(6, 1, 4, 7, 3, 9, 8, 2, 5),$$

the red and blue subsequences form a pair of twins of length 3, both similar to permutation $(1, 3, 2)$.

For a permutation π, let $t(\pi)$ denote the maximum integer k such that π contains twins of length k each. Further, let $t^{\text{perm}}(n)$ be the minimum of $t(\pi)$ over all permutations of $[n]$, called also *n-permutations*. By the first moment method Gawron [6] proved that $t^{\text{perm}}(n) \leqslant cn^{2/3}$ for some constant $c > 0$.

As for a lower bound, notice that by the Erdős-Szekeres Theorem, we have $t^{\text{perm}}(n) \geqslant \left\lfloor \frac{1}{2} n^{1/2} \right\rfloor$. This bound was substantially improved by Bukh and Rudenko [2]

Theorem 4 (Bukh and Rudenko [2]). *For all n, $t^{\mathrm{perm}}(n) \geqslant \frac{1}{8}n^{3/5}$.*

We call an ordered matching M on the set $[2n]$ *permutational* if the left end of each edge of M lies in the set $[n]$. In the double occurrence word notation such a matching can be written as $M = A_1A_2\dots A_nA_{i_1}A_{i_2}\dots A_{i_n}$, where $\pi_M = (i_1, i_2, \dots, i_n)$ is a permutation of $[n]$ (see Fig. 8).

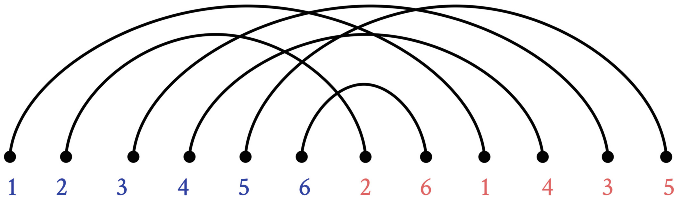

Fig. 8. The permutational matching that corresponds to the $(2, 6, 1, 4, 3, 5)$ permutation.

Clearly there are only $n!$ permutational matchings, nevertheless the connection to permutations turned out to be quite fruitful. Indeed, it is not hard to see that ordered twins in a permutational matching M are in one-to-one correspondence with twins in π_M. Thus, we have $t(M) = t(\pi_M)$ and, consequently, $t^{\mathrm{match}}(n) \leq t^{\mathrm{perm}}(n)$. In particular, by the above mentioned result of Gawron, it follows that $t^{\mathrm{match}}(n) = O(n^{2/3})$.

More subtle is the opposite relation.

Proposition 1. *For all $1 \leq m \leq n$, where $n - m$ is even,*

$$t^{\mathrm{match}}(n) \geq \min\left\{t^{\mathrm{perm}}(m),\ 2t^{\mathrm{match}}\left(\frac{n-m+2}{2}\right)\right\}.$$

Proof. Let M be a matching on $[2n]$. Split the set of vertices of M into two halves, $A = [n]$ and $B = [n+1, 2n]$ and let $M(A, B)$ denote the set of edges of M with one end in A and the other end in B. Note that $M' := M(A, B)$ is a permutational matching. We distinguish two cases. If $|M'| \geqslant m$, then

$$t(M) \geq t(M') = t(\pi_{M'}) \geq t^{\mathrm{perm}}(|M'|) \geq t^{\mathrm{perm}}(m).$$

If, on the other hand, $e(A, B) \leq m - 2$, then we have sub-matchings M_A and M_B of M of size at least $(n - m + 2)/2$ in sets, respectively, A and B. Thus, in this case, by concatenation,

$$t(M) \geq t(M_A) + t(M_B) \geq 2t^{\mathrm{match}}\left(\frac{n-m+2}{2}\right).$$

□

Proposition 1 allows, under some mild conditions, to „carry over" any lower bound on $t^{\mathrm{perm}}(n)$ to one on $t^{\mathrm{match}}(n)$.

Lemma 2. *If for some* $0 < \alpha, \beta < 1$*, we have* $t^{\text{perm}}(n) \geq \beta n^{\alpha}$ *for all* $n \geq 1$*, then* $t^{\text{match}}(n) \geq \beta(\gamma n)^{\alpha}$ *for any* $0 < \gamma \leq \min\{1 - 2^{1-1/\alpha}, 1/4\}$ *and all* $n \geq 1$*.*

Proof. Assume that for some $0 < \alpha, \beta < 1$, we have $t_r^{\text{perm}}(n) \geq \beta n^{\alpha}$ for all $n \geq 1$ and let $0 < \gamma \leq \min\{1 - 2^{1-1/\alpha}, 1/4\}$ be given. We will prove that $t^{\text{match}}(n) \geq \beta(\gamma n)^{\alpha}$ by induction on n. For $n \leqslant \frac{1}{\gamma}\left(\frac{1}{\beta}\right)^{1/\alpha}$ the claimed bound is at most 1, so it is trivially true. Assume that $n \geq \frac{1}{\gamma}\left(\frac{1}{\beta}\right)^{1/\alpha}$ and that $t^{\text{match}}(n') \geq \beta(\gamma n)^{\alpha}$ for all $n' < n$. Let $n_\gamma \in \{\lceil \gamma n \rceil, \lceil \gamma n \rceil + 1\}$ have the same parity as n. Then, by Proposition 1 with $m = n_\gamma$,

$$t^{\text{match}}(n) \geq \min\left\{t^{\text{perm}}(n_\gamma), 2t^{\text{match}}\left(\frac{n - n_\gamma + 2}{2}\right)\right\}.$$

By the assumption of the lemma, $t^{\text{perm}}(n_\gamma) \geq \beta n_\gamma^{\alpha} \geq \beta(\gamma n)^{\alpha}$. Since $\gamma \leq 1/4$ and so, $n \geq 4$, we have $(n - n_\gamma + 2)/2 \leq n - 1$. Hence, by the induction assumption, also

$$2t^{\text{match}}\left(\frac{n - n_\gamma + 2}{2}\right) \geq 2\beta\left(\gamma\frac{n - n_\gamma + 2}{2}\right)^{\alpha} \geq 2\beta\left(\gamma n\frac{1-\gamma}{2}\right)^{\alpha} \geq \beta(\gamma n)^{\alpha}$$

where the last inequality follows by the assumption on γ. $\square$

In particular, Theorem 4 and Lemma 2 with $\beta = 1/8$, $\alpha = 3/5$, and $\gamma = 1/4$ imply immediately the following result.

Corollary 1. *For every* n*,* $t^{\text{match}}(n) \geq \frac{1}{8}\left(\frac{n}{4}\right)^{3/5}$*.*

Moreover, any future improvement of the bound in Theorem 4 would automatically yield a corresponding improvement of the lower bound on $t^{\text{match}}(n)$.

As for an upper bound, we already mentioned that $t^{\text{match}}(n) = O\left(n^{2/3}\right)$. This means that for each n there is a matching M of size n with $t(M) \leq cn^{2/3}$, where $c > 0$ is a fixed constant. In fact, this holds for *almost all* M.

Proposition 2. *A.a.s.* $t(\mathbb{RM}_n) = O\left(n^{2/3}\right)$*.*

Proof. Consider a random (ordered) matching $\mathbb{RM}_n$. The expected number of twins of size k in $\mathbb{RM}_n$ is

$$\frac{1}{2}\binom{2n}{2k, 2k, 2n-4k}\frac{\alpha_k \cdot 1 \cdot \alpha_{n-2k}}{\alpha_n} = \frac{2^k n!}{2(2k)!k!(n-2k)!} < \left(\frac{e^3 n^2}{2k^3}\right)^k,$$

which tends to 0 with $n \to \infty$ if $k \geq cn^{2/3}$, for any $c > e2^{-1/3}$. This implies that a.a.s. there are no twins of size at least $cn^{2/3}$ in $\mathbb{RM}_n$. $\square$

4 Final Remarks

Proposition 2 asserts that a.a.s. $t(\mathbb{RM}_n) = O\left(n^{2/3}\right)$. In the journal version of this extended abstract we intend to prove the matching lower bound: a.a.s. $t(\mathbb{RM}_n) = \Omega\left(n^{2/3}\right)$. The real challenge, however, would be to prove (or disprove) that the bound holds for *all* matchings of size n.

Conjecture 1. For each n there is a matching M of size n with $t(M) \geq cn^{2/3}$, where $c > 0$ is a fixed constant. Consequently, $t^{\text{match}}(n) = \Theta\left(n^{2/3}\right)$.

The same statement is conjectured for twins in permutations (see [4]). By our results from Sect. 3, we know that both conjectures are actually equivalent.

In a similar way twins may be defined and studied in general ordered graphs.

Problem 1. How large twins must occur in every ordered graph with n edges?

For *unordered* graphs there is a result of Lee, Loh, and Sudakov [9] giving an asymptotically exact answer of order $\Theta(n \log n)^{2/3}$. It would be nice to have an analogue of this result for ordered graphs.

Finally, it seems natural to look for Erdős-Szekeres type results like Theorem 1 for more general structures. One possible direction to pursue is to consider, for some fixed $k \geqslant 3$, ordered *k-uniform* matchings. In full analogy with graph ordered matchings ($k = 2$), these structures correspond to *k-occurrence words*, in which every letter appears exactly k times. For instance, for $k = 3$ there are exactly $\frac{1}{2}\binom{6}{3} = 10$ ways two triples AAA and BBB can intertwine which, somewhat surprisingly, give rise to 9 canonical structures, analogous to lines, stacks, and waves in the graph case. In fact, they correspond to different pairs of the three graph structures. Using this correspondence, in the journal version we intend to prove that every 3-occurrence word of length $3n$ contains one of these 9 structures of size $\Omega\left(n^{1/9}\right)$. We suspect that similar phenomena hold for each $k \geq 4$ or even for words in which the occurrences of particular letters may vary.

Acknowledgements. We would like to thank all four anonymous referees for a careful reading of the manuscript and suggesting a number of editorial improvements.

References

1. Axenovich, M., Person, Y., Puzynina, S.: A regularity lemma and twins in words. J. Combin. Theor. Ser. A **120**(4), 733–743 (2013)
2. Bukh, B., Rudenko, O.: Order-isomorphic twins in permutations. SIAM J. Discret. Math. **34**(3), 1620–1622 (2020)
3. Conlon, D., Fox, J., Lee, C., Sudakov, B.: Ordered Ramsey numbers. J. Combin. Theor. Ser. B **122**, 353–383 (2017)
4. Dudek, A., Grytczuk, J., Ruciński, A.: Variations on twins in permutations. Electron. J. Combin. 28(3), Paper No. 3.19, 18 (2021)
5. Erdős, P., Szekeres, G.: A combinatorial problem in geometry. Compos. Math. **2**, 463–470 (1935)

6. Gawron, M.: Izomorficzne podstruktury w słowach i permutacjach, Master Thesis (2014)
7. Jelínek, V.: Dyck paths and pattern-avoiding matchings. Eur. J. Combin. **28**(1), 202–213 (2007)
8. Kasraoui, A., Zeng, J.: Distribution of crossings, nestings and alignments of two edges in matchings and partitions. Electron. J. Combin. 13(1), Research Paper 33, 12 (2006)
9. Lee, C., Loh, P.S., Sudakov, B.: Self-similarity of graphs. SIAM J. Discret. Math. **27**(2), 959–972 (2013)
10. Stanley, R.P.: Increasing and decreasing subsequences and their variants. In: International Congress of Mathematicians, Vol. I, pp. 545–579. European Mathematical Society, Zürich (2007)
11. Tardos, G.: Extremal theory of ordered graphs. In: Proceedings of the International Congress of Mathematicians-Rio de Janeiro 2018, Vol. IV, invited lectures, pp. 3235–3243. World Scientific Publishing, Hackensack, NJ (2018)

Author Queries

Chapter 33

Query Refs.	Details Required	Author's response
AQ1	This is to inform you that corresponding author has been identified as per the information available in the Copyright form.	