Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Journal of Combinatorial Theory, Series A 107 (2004) 153–160
Journal of Combinatorial Theory Series A
http://www.elsevier.com/locate/jcta

Note

Excluded permutation matrices and the Stanley–Wilf conjecture

Adam Marcus[1] and Gábor Tardos[2]

Alfréd Rényi Institute, 1364 Budapest Pf.127, Hungary

Received 18 December 2003

Abstract

This paper examines the extremal problem of how many 1-entries an $n \times n$ 0–1 matrix can have that avoids a certain fixed submatrix P. For any permutation matrix P we prove a linear bound, settling a conjecture of Zoltán Füredi and Péter Hajnal (Discrete Math. 103(1992) 233). Due to the work of Martin Klazar (D. Krob, A.A. Mikhalev, A.V. Mikhalev (Eds.), Formal Power Series and Algebraics Combinatorics, Springer, Berlin, 2000, pp. 250–255), this also settles the conjecture of Stanley and Wilf on the number of n-permutations avoiding a fixed permutation and a related conjecture of Alon and Friedgut (J. Combin Theory Ser A 89(2000) 133).

Keywords: Pattern avoidance; Extremal problems; Stanley-Wilf conjecture; Forbidden submatrices

1. Introduction

This paper settles three related conjectures concerning pattern avoidance. To state the conjectures we define the term "avoiding" in several contexts.

Definition. Let A and P be 0–1 matrices. We say that A *contains* the $k \times l$ matrix $P = (p_{ij})$ if there exists a $k \times l$ submatrix $D = (d_{ij})$ of A with $d_{ij} = 1$ whenever $p_{ij} = 1$. Otherwise we say that A *avoids* P. Notice that we can delete rows and columns of A

E-mail addresses: adam@math.gatech.edu (A. Marcus), tardos@renyi.hu (G. Tardos).

[1] Visiting Researcher, Alfréd Rényi Institute, 1364 Budapest, Pf.127, Hungary, on leave from the Department of Mathematics (ACO), Georgia Institute of Technology, Atlanta, GA 30332-0160. This research was made possible due to funding by the Fulbright Program in Hungary.

[2] Partially supported by the Hungarian National Research Fund OTKA T029255 and T046234.

0097-3165/$ - see front matter
doi:10.1016/j.jcta.2004.04.002

to obtain the submatrix D but we cannot permute the remaining rows and columns. If A contains P we identify the 1-entries of the matrix A corresponding to the entries d_{ij} of D with $p_{ij} = 1$ and say that these entries of A *represent* P.

Let $[n] = \{1, 2, \dots, n\}$. A permutation of $[n]$ is called an n-permutation. We say that an n-permutation σ *contains* a k-permutation π if there exist integers $1 \leqslant x_1 < x_2 < \cdots < x_k \leqslant n$ such that for $1 \leqslant i, j \leqslant k$ we have

$$\sigma(x_i) < \sigma(x_j) \text{ if and only if } \pi(i) < \pi(j).$$

Otherwise, we say that σ *avoids* π.

The set of finite sequences (words) over $[n]$ is denoted by $[n]^*$. A sequence $a = a_1 a_2 \dots a_l \in [n]^*$ *contains* the sequence $b = b_1 b_2 \dots b_k \in [m]^*$ if there exist indices $1 \leqslant x_1 < x_2 < \cdots < x_k \leqslant l$ such that for $1 \leqslant i, j \leqslant k$ we have

$$a_{x_i} < a_{x_j} \text{ if and only if } b_i < b_j.$$

Otherwise, we say that *a avoids b*.

Note that these notions are intimately related. It is easy to see that the word $\sigma(1)\sigma(2)\dots\sigma(n)$ contains the word $\pi(1)\pi(2)\dots\pi(k)$ if and only if the n-permutation σ contains the k-permutation π, which happens if and only if the permutation matrix of σ contains the permutation matrix of π.

Definition. For a 0–1 matrix P let $f(n, P)$ be the maximum number of 1-entries in an $n \times n$ 0–1 matrix avoiding P.

For a permutation π let $S_n(\pi)$ be the number of n-permutations avoiding π.

A sequence $a_1 a_2 \dots a_n$ is considered *k-sparse* if $i < j$, $a_i = a_j$ implies $j - i \geqslant k$. For a sequence $b \in [m]^*$ and $k \geqslant m$ let $l_k(b, n)$ be the maximum length of a k-sparse word in $[n]^*$ avoiding b.

The main result of this paper is the following theorem proving the Füredi–Hajnal conjecture (originally posed in [8]).

Theorem 1. *For all permutation matrices P we have $f(n, P) = O(n)$.*

By a result of Martin Klazar [12] (which we reproduce in Section 3) the above result proves the Stanley–Wilf conjecture (stated here as Corollary 2) and also the Alon–Friedgut conjecture (stated here as Corollary 3), which was originally posed in [1]. The Stanley–Wilf conjecture was formulated by Richard Stanley and Herbert Wilf around 1992 but it is hard to find an exact reference. An even earlier source is the Ph.D. thesis of Julian West [17] of 1990 where he asks about the growth rate of $S_n(\pi)$. His Question 3.4.3 is more specific; he asks if $S_n(\pi)$ and $S_n(\pi')$ are asymptotically equal for k-permutations π and π'. Miklós Bóna [4] showed that this conjecture was too strong, however, by finding 4-permutations π and π' with $S_n(\pi)$ and $S_n(\pi')$ displaying different growth rates. Nevertheless, it shows a direct interest in asymptotic enumerations of this kind.

Corollary 2. *For all permutations* π *there exists a constant* $c = c_\pi$ *such that* $S_n(\pi) \leqslant c^n$.

Corollary 3. *For a* k*-permutation* σ *and the word* $a = \sigma(1)\sigma(2)\ldots\sigma(k)$ *we have* $l_k(a, n) = O(n)$.

Several special cases of the above conjectures have already been established. Bóna [5] proved the Stanley–Wilf conjecture for *layered* permutations π, that is, for permutations consisting of an arbitrary number of increasing blocks with all elements of a block smaller than the elements of the previous block. Alon and Friedgut [1] proved the conjecture for permutations consisting of an increasing sequence followed by a decreasing one or vice versa. Approximate versions of these conjectures have also been established. Using a result of Klazar [9] on generalized Davenport–Schinzel sequences Alon and Friedgut [1] showed approximate versions of their own conjecture and the Stanley–Wilf conjecture where the linear and exponential bounds were replaced by $O(n\gamma(n))$ and $2^{O(n\gamma(n))}$, respectively, with an extremely slow growing function γ related to the inverse Ackermann function.

In Section 2 we give a surprisingly simple and straightforward proof of Theorem 1. For the reader's convenience we reproduce Klazar's argument on how this result implies the Stanley–Wilf conjecture in Section 3. In Section 4 we discuss further consequences of our main theorem as well as related problems that are still open.

2. Proof of the Füredi–Hajnal conjecture

Theorem 1 is proved by establishing a linear recursion for $f(n, P)$ in Lemma 7, that in turn is based on three rather simple lemmas. We partition the larger matrix into blocks. This idea appears in several related papers, e.g. in [12], but we use larger blocks than were previously considered.

Throughout these lemmas, we let P be a fixed $k \times k$ permutation matrix and A be an $n \times n$ matrix with $f(n, P)$ 1-entries which avoids P. We assume k^2 divides n. We define S_{ij} to be the square submatrix of A consisting of the entries $a_{i'j'}$ with $i' \in [k^2(i-1)+1, k^2 i]$, $j' \in [k^2(j-1)+1, k^2 j]$. We let $B = (b_{ij})$ be the $(n/k^2) \times (n/k^2)$ 0–1 matrix with $b_{ij} = 0$ if and only if all entries of S_{ij} are zero. We say that a block is *wide* (respectively, *tall*) if it contains 1-entries in at least k different columns (respectively rows).

Lemma 4. *B avoids P.*

Proof. Assume not and consider the k 1-entries of B representing P. Choose an arbitrary 1-entry from the k corresponding blocks of A. They represent P, contradicting the fact that A avoids P. $\square$

Lemma 5. *Consider the set (column) of blocks* $C_j = \{S_{ij} : i = 1, \ldots, \frac{n}{k^2}\}$. *The number of blocks in* C_j *that are wide is less than* $k\binom{k^2}{k}$.

Proof. Assume not. By the pigeonhole principle, there exist k blocks in C_j that have a 1-entry in the same columns $c_1 < c_2 < \cdots < c_k$. Let $S_{d_1 j}, \ldots, S_{d_k j}$ be these blocks with $1 \leqslant d_1 < d_2 < \cdots < d_k \leqslant n/k^2$. For each 1-entry p_{rs}, pick any 1-entry in column c_s of $S_{d_r j}$. These entries of A represent P, a contradiction. □

Lemma 6. *Consider the set (row) of blocks $R_i = \{S_{ij} : j = 1, \ldots, \frac{n}{k^2}\}$. The number of blocks in R_i that are tall is less than $k\binom{k^2}{k}$.*

Proof. The same proof applies as for Lemma 5. □

With these tools, the main lemma follows:

Lemma 7. *For a $k \times k$ permutation matrix P and n divisible by k^2 we have*

$$f(n, P) \leqslant (k-1)^2 f\left(\frac{n}{k^2}, P\right) + 2k^3 \binom{k^2}{k} n.$$

Proof. We consider three types of blocks:

- $X_1 = \{\text{blocks that are wide}\}$.

  $|X_1| \leqslant \frac{n}{k^2} k \binom{k^2}{k}$ by Lemma 5.

- $X_2 = \{\text{blocks that are tall}\}$.

  $|X_2| \leqslant \frac{n}{k^2} k \binom{k^2}{k}$ by Lemma 6.

- $X_3 = \{\text{nonempty blocks that are neither wide nor tall}\}$.

  $|X_3| \leqslant f(\frac{n}{k^2}, P)$ by Lemma 4.

This includes all of the nonempty blocks. We bound $f(n, P)$, the number of ones in A, by summing estimates of the number of ones in these three categories of blocks. Any block contains at most k^4 1-entries and a block of X_3 contains at most $(k-1)^2$ 1-entries. Thus,

$$\begin{aligned} f(n, P) &\leqslant k^4 |X_1| + k^4 |X_2| + (k-1)^2 |X_3| \\ &\leqslant 2k^3 \binom{k^2}{k} n + (k-1)^2 f\left(\frac{n}{k^2}, P\right). \qquad \square \end{aligned}$$

Solving the above linear recursion gives the following Theorem and also Theorem 1. We did not optimize for the constant factor here.

Theorem 8. *For a $k \times k$ permutation matrix P we have*

$$f(n,P) \leqslant 2k^4 \binom{k^2}{k} n.$$

Proof. We proceed by induction on n. The base cases (when $n \leqslant k^2$) are trivial. Now assume the hypothesis to be true for all $n < n_0$ and consider the case $n = n_0$. We let n' be the largest integer less than or equal to n which is divisible by k^2. Then by Lemma 7, we have:

$$\begin{aligned} f(n,P) &\leqslant f(n',P) + 2k^2 n \\ &\leqslant (k-1)^2 f\left(\frac{n'}{k^2}, P\right) + 2k^3 \binom{k^2}{k} n' + 2k^2 n \\ &\leqslant (k-1)^2 \left[2k^4 \binom{k^2}{k} \frac{n'}{k^2}\right] + 2k^3 \binom{k^2}{k} n' + 2k^2 n \\ &\leqslant 2k^2((k-1)^2 + k + 1) \binom{k^2}{k} n \\ &\leqslant 2k^4 \binom{k^2}{k} n \end{aligned}$$

where the last inequality is true for all $k \geqslant 2$. □

3. Deduction of the Stanley–Wilf conjecture

For the reader's convenience (and to show the similarities of the two proofs) we sketch here Klazar's argument [13] that the Füredi–Hajnal conjecture implies the Stanley–Wilf conjecture.

Definition. For a 0–1 matrix P let $T_n(P)$ be the set of $n \times n$ matrices which avoid P.

As we noted in Section 1, a permutation σ avoids another permutation π if and only if the permutation matrix corresponding to σ avoids the permutation matrix corresponding to π. Thus if P is the permutation matrix of the permutation π, then $T_n(P)$ contains the permutation matrices of all n-permutations avoiding π. In particular we have $|T_n(p)| \geqslant S_n(\pi)$.

Assuming the Füredi–Hajnal conjecture, Klazar proves the following statement, which in turn implies Corollary 2:

Theorem 9. *For any permutation matrix P there exists a constant $c = c_P$ such that $|T_n(P)| \leqslant c^n$.*

Proof. Using $f(n,P) = O(n)$ the statement of the theorem follows from the following simple recursion:

$$|T_{2n}(P)| \leqslant |T_n(P)| 15^{f(n,P)}.$$

To prove the recursion we map $T_{2n}(P)$ to $T_n(P)$ by partitioning any matrix $A \in T_{2n}(P)$ into 2×2 blocks and replacing each all-zero block by a 0-entry and all other blocks by 1-entries. As we saw in Lemma 4 the resulting $n \times n$ matrix B avoids P. Any matrix $B \in T_n(P)$ is the image of at most 15^w matrices of $T_{2n}(P)$ under this mapping where w is the number of 1-entries in B. Here $w \leqslant f(n,P)$ so the recursion and the Theorem follow. □

The reduction also provides a nice characterization in the theory of excluded matrices:

Corollary 10. *For any* 0–1 *matrix* P, *we have* $\log(|T_n(P)|) = O(n)$ *if and only if* P *has at most a single* 1-*entry in each row and column.*

Proof. The matrices in the characterization are the submatrices of permutation matrices. For these matrices $\log(|T_n(P)|) = O(n)$ follows from Theorem 9. For other matrices P, $T_n(P)$ contains all of the $n \times n$ permutation matrices (a total of $n!$), so $\log(|T_n(P)|) = \Omega(n \log n)$. □

4. Generalizations and open problems

The problem of estimating the extremal function $f(n,P)$ for 0–1 matrices P was considered first for some special patterns P in [3,7]. Later [8] systematically treated all patterns P with at most four 1-entries and established the order of magnitude of $f(n,P)$ for almost all of them. For the missing few such patterns see [16], where certain *sets of avoided patterns* are also considered. We restate the following question of [8].

Problem 1 (Füredi–Hajnal [8]). *What are the matrices* P *with* $f(n,P) = O(n)$?

This problem seems to be much harder than the corresponding enumerative problem settled by Corollary 10. A characterization will probably be based on excluded submatrices. We therefore define P to be *minimally nonlinear* if $f(n,P)$ is nonlinear but replacing any 1-entry in P with 0 yields a pattern P' with $f(n,P') = O(n)$. We can also exclude patterns containing an empty row or column. As any 0–1 matrix with three 1-entries has a linear extremal function any pattern P containing four 1-entries with $f(n,P)$ nonlinear is minimally nonlinear. These patterns are all known, but no other minimally nonlinear patterns are known. We pose the following problem:

Problem 2. *Find any minimally nonlinear patterns P with more than four 1-entries.*

Consider bipartite graphs with a fixed bipartition and separate linear orders on both sets of vertices. Notice that the ordered bipartite graph G_1 contains another ordered bipartite graph G_2 exactly if, for their (bipartite) adjacency matrices M_1 and M_2, M_1 contains M_2. Thus one can interpret $f(n,P)$ as the maximum number of edges of an ordered bipartite graph on $n+n$ vertices which does not contain a certain (ordered bipartite) subgraph. This interpretation makes the study of $f(n,P)$ the ordered bipartite version of the Turán-type extremal graph theory. In this interpretation our theorem states that (ordered bipartite) matchings have linear extremal functions. One can also study general graphs on an ordered vertex set. See more on the relation between ordered graphs and excluded matrices in [14]. Peter Brass et al. [6] studied the extremal theory of graphs with a *cyclically ordered* vertex set. In the unordered case the problem of finding the minimally nonlinear (bipartite) graphs is easy: they are the (even) cycles. The ordered problem seems to be far more complex. Several works of Klazar [10,11,13] consider the generalization to ordered hypergraphs. Among other results, one can find far reaching hypergraph consequences of the Füredi–Hajnal conjecture there. For the sake of brevity, we do not state these interesting enumerative and extremal hypergraph results here.

All papers on the Stanley–Wilf conjecture mention the original and stronger form of the conjecture: Is it true that for any permutation σ the numbers $(S_n(\sigma))^{1/n}$ tend to a finite limit c_σ as n goes to infinity? By the result of Richard Arratia [2] the two forms of the Stanley–Wilf conjecture are equivalent, so now we have that $c_\sigma < \infty$ exists. Working out the bounds for a k-permutation σ in Theorem 8 one finds an explicit bound $c_\sigma \leqslant 15^{2k^4\binom{k^2}{k}}$. This doubly exponential bound is very far from the one conjectured by Richard Arratia [2].

Problem 3 (Arratia [2]). $c_\sigma \leqslant (k-1)^2$.

Note that Regev's asymptotic formula for $S_n(Id_k)$ [15] implies $c_{Id_k} = (k-1)^2$ for the identity k-permutation Id_k.

With respect to the Alon–Friedgut conjecture we can ask the following question:

Problem 4. *What are the words* $a \in [k]^*$ *with* $l_k(a,n) = O(n)$?

This problem turns out to be an interesting special case of Problem 1. For a word $a = a_1 a_2 \ldots a_l \in [k]^*$ we have $l_k(a,n) = O(n)$ if and only if $f(n,P_a) = O(n)$ for the $k \times l$ 0–1 matrix $P = (p_{ij})$ where $p_{ij} = 1$ exactly if $a_j = i$. This equivalence can be proved along the lines of [12]. Notice that $f(n,P_a)$ is nonlinear for many words a, e.g. for $a = 1212$ or even for $a = 1213$, but it is linear for seemingly similar other words like $a = 1312$. One may try to prove linearity of $f(n,P_a)$ for certain words a using techniques similar to those presented in Section 2, but Lemma 4 holds only for (submatrices of) permutation matrices.

Acknowledgments

We would like to thank Martin Klazar and Miklós Bóna for the great help they provided in writing this paper. Among other things Martin Klazar showed us how our result gives the characterization stated in Corollary 10 and Miklós Bóna showed us the first written reference to the growth rate of $S_n(\pi)$ in [17].

References

[1] N. Alon, E. Friedgut, On the number of permutations avoiding a given pattern, J. Combin. Theory Ser. A 89 (2000) 133–140.

[2] R. Arratia, On the Stanley–Wilf conjecture for the number of permutations avoiding a given pattern, Electron. J. Combin. 6(1) 1999 1–4, http://www.combinatorics.org/Volume_6/Abstracts/v6i1n1.html.

[3] D. Bienstock, E. Győri, An extremal problem on sparse 0-1 matrices, SIAM J. Discrete Math. 4 (1) (1991) 17–27.

[4] M. Bóna, Exact and asymptotic enumeration of permutations with subsequence conditions, Ph.D. Thesis, MIT, Cambridge, MA, 1997.

[5] M. Bóna, The solution of a conjecture of Stanley and Wilf for all layered patterns, J. Combin. Theory Ser. A 85 (1999) 96–104.

[6] P. Brass, G. Károlyi, P. Valtr, A Turán-type extremal theory of convex geometric graphs, in: B. Aronov, S. Basu, J. Pach, M. Sharir (Eds.), Discrete and Computational Geometry—The Goodman–Pollack Festschrift, Springer, Berlin, 2003, pp. 275–300.

[7] Z. Füredi, The maximum number of unit distances in a convex n-gon, J. Combin. Theory Ser. A 55 (2) (1992) 316–320.

[8] Z. Füredi, P. Hajnal, Davenport–Schinzel theory of matrices, Discrete Math. 103 (1992) 231–251.

[9] M. Klazar, A general upper bound in extremal theory of sequences, Comment. Math. Univ. Carolin. 33 (1992) 737–746.

[10] M. Klazar, Counting pattern-free set partitions I: a generalization of Stirling numbers of the second kind, Europ. J. Combin. 21 (2000) 367–378.

[11] M. Klazar, Counting pattern-free set partitions II: noncrossing and other hypergraphs, Electron. J. Combin. R34 (7) (2000) 25.

[12] M. Klazar, The Füredi–Hajnal conjecture implies the Stanley–Wilf conjecture, in: D. Krob, A.A. Mikhalev, A.V. Mikhalev (Eds.), Formal Power Series and Algebraic Combinatorics, Springer, Berlin, 2000, pp. 250–255.

[13] M. Klazar, Extremal problems for ordered (hyper)graphs: applications of Davenport–Schinzel sequences, Europ. J. Combin. 25 (2004) 125–140.

[14] J. Pach, G. Tardos, Forbidden paths and cycles in ordered graphs and matrices, manuscript, 2004, http://www.renyi.hu/~tardos/papers.html.

[15] A. Regev, Asymptotic values for degrees associated with strips of Young diagrams, Adv. Math. 41 (1981) 115–136.

[16] G. Tardos, On 0–1 matrices and small excluded submatrices, manuscript, 2004, http://www.renyi.hu/~tardos/papers.html.

[17] J. West, Permutations with restricted subsequences and stack-sortable permutations, Ph.D. Thesis, MIT, Cambridge, MA, 1990.