

Graph Parameters and Ramsey Theory

Vadim Lozin^{1,2} (\boxtimes)

¹ Mathematics Institute, University of Warwick, Coventry CV4 7AL, UK V.Lozin@warwick.ac.uk

 $^{2}\,$ Lobachevsky State University of Nizhni Novgorod, Nizhny Novgorod, Russia

Abstract. Ramsey's Theorem tells us that there are exactly two minimal hereditary classes containing graphs with arbitrarily many vertices: the class of complete graphs and the class of edgeless graphs. In other words, Ramsey's Theorem characterizes the graph vertex number in terms of minimal hereditary classes where this parameter is unbounded. In the present paper, we show that a similar Ramsey-type characterization is possible for a number of other graph parameters, including neighbourhood diversity and VC-dimension.

1 Introduction

In 1930, a 26 years old British mathematician Frank Ramsey proved the following theorem, known nowadays as Ramsey's Theorem.

Theorem 1 [12]. For any positive integers k, r, p, there exists a minimum positive integer F = F(k, r, p) such that if the k-subsets of an F-set are colored with r colors, then there is a monochromatic p-set, i.e. a p-set all of whose k-subsets have the same color.

It is not difficult to see that with k = 1 this theorem coincides with the Pigeonhole Principle. For k = 2, the theorem admits a nice interpretation in the terminology of graph theory, since coloring 2-subsets can be viewed as coloring the edges of a complete graph. In the case of r = 2 colors, the graph-theoretic interpretation of Ramsey's Theorem can be further rephrased as follows.

Theorem 2. For any positive integer p, there is a minimum positive integer R = R(p) such that every graph with at least R vertices has either a clique of size p or an independent set of size p.

It is not difficult to see that Theorem 2 is equivalent to the following statement.

Theorem 3. The class of complete graphs and the class of edgeless graphs are the only two minimal infinite hereditary classes of graphs.

Theorem 3 characterizes the family of hereditary classes containing graphs with a bounded number of vertices in terms of minimal "forbidden" elements, i.e. minimal hereditary classes where the vertex number is unbounded. Is it possible to find a similar characterization for other graph parameters?

The purpose of this paper is to show that Ramsey's theorem can be used to find minimal "forbidden" classes for various other parameters. For instance, directly from Ramsey's Theorem it follows that

- the class of complete graphs and the class of stars (and all their induced subgraphs) are the only two minimal hereditary classes of graphs of unbounded vertex degree.
- the class of complete graphs and the class of complete bipartite graphs are the only two minimal hereditary classes of graphs of unbounded biclique number (the size a maximum complete bipartite subgraph with equal parts).

In the subsequent sections, we show that a similar Ramsey-type characterization is possible for a number of other graph parameters. In the rest of the present section, we introduce basic definitions and notations used in the paper.

We consider only simple undirected graphs without loops and multiple edges and denote the vertex set and the edge set of a graph G by V(G) and E(G), respectively. If v is a vertex of G, then N(v) is its neighbourhood, i.e. the set of vertices of G adjacent to v. The closed neighbourhood of v is defined and is denoted as $N[v] = N(v) \cup \{v\}$. The degree of v is |N(v)|.

In a graph, an *independent set* is a subset of vertices no two of which are adjacent, a *clique* is a subset of vertices every two of which are adjacent, and a *matching* is a subset of edges no two which share a vertex.

For a graph G, we denote by \overline{G} the complement of G. Similarly, for a class \mathcal{X} of graphs, we denote by $\overline{\mathcal{X}}$ the class of complements of graphs in \mathcal{X} .

Given a graph G and a subset $U \subseteq V(G)$, we denote by G[U] the subgraph of G induced by U, i.e. the subgraph obtained from G by deleting all the vertices not in U. We say that a graph G contains a graph H as an induced subgraph if His isomorphic to an induced subgraph of G. A graph G is said to be *n*-universal for a class of graphs \mathcal{X} if G contains all *n*-vertex graphs from \mathcal{X} as induced subgraphs.

A class \mathcal{X} of graphs is *hereditary* if it is closed under taking induced subgraphs, i.e. if $G \in \mathcal{X}$ implies $H \in \mathcal{X}$ for every graph H contained in G as an induced subgraph. Two hereditary classes of particular interest in this paper are split graphs and bipartite graphs.

A graph G is a *split* graph if V(G) can be partitioned into an independent set and a clique, and G is *bipartite* if V(G) can be partitioned into at most two independent sets. A bipartite graph G given together with a bipartition of its vertices into independent sets A and B will be denoted G = (A, B, E), in which case we will say that A and B are the color classes of G. If every vertex of A is adjacent to every vertex of B, then G = (A, B, E) is *complete bipartite*. The *bipartite complement* of G = (A, B, E) is the bipartite graph G' = (A, B, E'), where $ab \in E'$ if and only if $ab \notin E$. Clearly, by creating a clique in one of the color classes of a bipartite graph, we transform it into a split graph, and vice versa.

2 Neighbourhood Diversity

The neighbourhood diversity of a graph was introduced in [7] and was used to develop fpt-algorithms for some difficult graph problems (see e.g. [5]). This parameter can be defined as follows.

Definition 1. Two vertices x and y are said to be similar if there is no vertex z distinguishing them i.e. if there is no vertex z adjacent to exactly one of x and y. Clearly, the similarity is an equivalence relation. The neighbourhood diversity of G is the number of similarity classes in G.

Before we provide a Ramsey-type characterization of the neighbourhood diversity, we introduce an auxiliary notion.

Definition 2. A skew matching in a graph G is a matching $\{x_1y_1, \ldots, x_qy_q\}$ such that y_i is not adjacent to x_j for all i < j. The complement of a skew matching is a sequence of pairs of vertices that create a skew matching in the complement of G.

Lemma 1. For any positive integer m, there exists a positive integer r = r(m) such that any bipartite graph G = (A, B, E) of neighbourhood diversity r contains either a skew matching of size m or its complement.

Proof. Define $r = 2^{2m}$ and let X be a set of pairwise non-similar vertices of size r/2 chosen from the same color class of G, say from A. Let y_1 be a vertex in B distinguishing the set X (i.e. y_1 has both a neighbour and a non-neighbour in X) and let us say that y_1 is big if the number of its neighbours in X is larger than the number of its non-neighbours, and small otherwise. If y_1 is small, we arbitrarily choose its neighbour in X, denote it by x_1 and remove all neighbours of y_1 from X. If y is big, we arbitrarily choose a non-neighbour of y_1 in X, denote it by x_1 and remove all non-neighbours of y_1 from X. Observe that y_1 does not distinguish the vertices in the updated set X.

We apply the above procedure to $X \ 2m - 1$ times and obtain in this way a sequence of 2m - 1 pairs $x_i y_i$. If m of these pairs contain small vertices y_i , then the respective pairs create a skew matching of size m. Otherwise, there is a set of m pairs containing big vertices y_i , in which case the respective pairs create the complement of a skew matching.

Now we turn to the neighbourhood diversity and start with the bipartite case. For this, we denote by

- \mathcal{M} the class of graphs of vertex degree at most 1. By M_n we denote an induced matching of size n, i.e. the unique up to isomorphism graph in the class \mathcal{M} with 2n vertices each of which has degree 1. Clearly, M_n is n-universal for graphs in \mathcal{M} .
- \mathcal{M}^{bc} the class of bipartite complements of graphs in \mathcal{M} . The bipartite complement of the graph M_n will be denoted M_n^{bc} . Clearly, M_n^{bc} is *n*-universal for graphs in \mathcal{M}^{bc} .

 \mathcal{Z} the class of *chain graphs*, i.e. bipartite graphs in which the neighbourhoods of the vertices in each part form a chain with respect to set-inclusion. By Z_n we denote a chain graph such that for each $i \in \{1, 2, ..., n\}$, each part of the graph contains exactly one vertex of degree *i*. Figure 1 represents the graph Z_n for n = 5. It is known [9] that Z_n is *n*-universal for graphs in \mathcal{Z} .

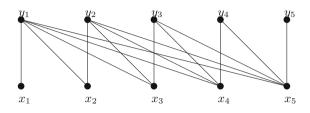


Fig. 1. The graph Z_5

Lemma 2. For any positive integer p, there exists a positive integer q = q(p) such that any bipartite graph G = (A, B, E) of neighbourhood diversity q contains either an induced M_p or an induced Z_p or an induced M_p^{bc} .

Proof. Let m = R(p+1) (where R is the Ramsey number defined in Theorem 2) and $q = 2^{2m}$. According to the proof of Lemma 1, G contains a skew matching of size m or its complement. If G contains a skew matching M, we color each pair (x_iy_i, x_jy_j) of edges of M (i < j) in two colors as follows:

- color 1 if x_i is not adjacent to y_j ,
- color 2 if x_i is adjacent to y_j .

By Ramsey's Theorem, M contains a monochromatic set M' of edges of size p + 1. If the color of each pair of edges in M' is

- 1. then M' is an induced matching of size p + 1,
- 2. then the vertices of M' induce a Z_{p+1} .

Analogously, in the case when G contains the complement of a skew matching, we find either an induced M_{p+1}^{bc} or an induced Z_p (observe that the bipartite complement of Z_{p+1} contains an induced Z_p).

Now we proceed to the general case and denote by

- \mathcal{M}^* the class of split graphs obtained from graphs in \mathcal{M} by creating a clique in one of the color classes. The graph obtained from M_n by creating clique in one its color classes will be denoted by M_n^* . Clearly, M_n^* is *n*-universal for graphs in \mathcal{M}^* .
- \mathcal{Z}^* the class of split graphs obtained from graphs in \mathcal{Z} by creating a clique in one of the color classes. This class is known in the literature as the class of *threshold graphs*. The graph obtained from Z_n by creating a clique in one of its color classes will be denoted Z_n^* . This graph is *n*-universal for threshold graphs [6].

Lemma 3. For any positive integer p, there exists a positive integer Q = Q(p)such that every graph G of neighbourhood diversity Q contains one of the following nine graphs as an induced subgraph: M_p , M_p^{bc} , Z_p , \overline{M}_p , \overline{M}_p^{bc} , \overline{Z}_p , M_p^* , \overline{M}_n^* , Z_n^* .

Proof. Let Q = R(q), where $q = 2^{2m}$ and m = R(R(p) + 1) (R is the Ramsey number). We choose one vertex from each similarity class of G and find in the chosen set a subset A of vertices that form an independent set or a clique of size $q = 2^{2m}$. Let us call the vertices of A white. We denote the remaining vertices of G by B and call them black. Let G' denote the bipartite subgraph of G formed by the edges between A and B. By the choice of A, all vertices of this set have pairwise different neighbourhoods in G'. Therefore, according to the proof of Lemma 2, G' contains a subgraph G'' inducing either M_n , or M_n^{bc} or Z_n with n = R(p). Among the n black vertices of G'', we can find a subset B' of vertices that form either a clique or an independent set of size p in the graph G. Then B' together with a subset of A of size p induce in G one of the nine graphs listed in the statement of the lemma.

Since the nine graphs of Lemma 3 are universal for their respective classes, we make the following conclusion.

Theorem 4. There exist exactly nine minimal hereditary classes of graphs of unbounded neighbourhood diversity: $\mathcal{M}, \mathcal{M}^{bc}, \mathcal{Z}, \overline{\mathcal{M}}, \overline{\mathcal{M}}^{bc}, \overline{\mathcal{Z}}, \mathcal{M}^*, \overline{\mathcal{M}}^*, \mathcal{Z}^*.$

3 VC-Dimension

A set system (X, S) consists of a set X and a family S of subsets of X. A subset $A \subseteq X$ is *shattered* if for every subset $B \subseteq A$ there is a set $C \in S$ such that $B = A \cap C$. The VC-dimension of (X, S) is the cardinality of a largest shattered subset of X.

The VC-dimension of a graph G = (V, E) was defined in [1] as the VCdimension of the set system (V, S), where S the family of closed neighbourhoods of vertices of G, i.e. $S = \{N[v] : v \in V(G)\}$. Let us denote the VC-dimension of G by vc[G].

In this section, we characterize VC-dimension by means of three minimal hereditary classes where this parameter is unbounded. To this end, we first redefine it in terms of open neighbourhoods as follows. Let vc(G) be the size of a largest set A of vertices of G such that for any subset $B \subseteq A$ there is a vertex v outside of A with $B = A \cap N(v)$. In other words, vc(G) is the size of a largest subset of vertices shattered by open neighbourhoods of vertices of G.

We start by showing that the two definitions are equivalent in the sense that they both are either bounded or unbounded in a hereditary class. To prove this, we introduce the following terminology. Let A be a set of vertices which is shattered by a collection of closed neighbourhoods. For a subset $B \subseteq A$ we will denote by v(B) the vertex whose neighbourhood intersect A at B. We will say that B is closed if v(B) belongs to B, and open otherwise.

Lemma 4. $vc(G) \le vc[G] \le vc(G)(vc(G) + 1) + 1$

Proof. The first inequality is obvious. To prove the second one, let A be a subset of V(G) of size vc[G] which is shattered by a collection of closed neighbourhoods. If A has no closed subsets, then vc[G] = vc(G). Otherwise, let B be a closed subset of A.

Assume first that |B| = 1. Then $B = \{v(B)\}$ and v(B) is isolated in G[A], i.e. it has no neighbours in A. Let C be the set of all such vertices, i.e. vertices each of which is a closed subset of A. By removing from A any vertex $x \in C$ we obtain a new set A and may assume that it has no closed subsets of size 1. Indeed, for any vertex $y \in C$ different from x, there must exist a vertex $y' \notin A$ such that $N(y') \cap A = \{x, y\}$ (since A is shattered). After the removal of x from A, we have $N(y') \cap A = \{y\}$ and hence $\{y\}$ is not a closed subset anymore. This discussion allows us to assume in what follows that A has no closed subsets of size 1, in which case we only need to show that $vc[G] \leq vc(G)(vc(G) + 1)$.

Assume now that B is a closed subset of A of size at least 2. Suppose that B-v(B) contains a closed subset C, i.e. $v(C) \in C$. Observe that v(C) is adjacent to v(B), as every vertex of B-v(B) is adjacent to v(B). But then $N[v(C)] \cap A$ contains v(B) contradicting the fact that $N[v(C)] \cap A = C$. This contradiction shows that every subset of B-v(B) is open, i.e. $|B-v(B)| \leq vc(G)$.

The above observation allows us to apply the following procedure: as long as A contains a closed subset B with at least two vertices, delete from A all vertices of B except for v(B). Denote the resulting set by A^* . Assume the procedure was applied p times and let B_1, \ldots, B_p be the closed subsets it was applied to. It is not difficult to see that the set $\{v(B_1), \ldots, v(B_p)\}$ has no closed subsets and hence its size cannot be large than vc(G), i.e. $p \leq vc(G)$. Combining, we conclude:

$$vc[G] = |A| \le |A^*| + \sum_{i=1}^p |B_i - v(B_i)| \le vc(G) + p \cdot vc(G) \le vc(G)(vc(G) + 1).$$

This lemma allows us to assume that if A is shattered, then there is a set C disjoint from A such that for any subset $B \subseteq A$ there is a vertex $v \in C$ with $B = A \cap N(v)$, in which case we will say that A is shattered by C, or C shatters A.

Let $Q_n = (A, B, E)$ be the bipartite graph with |A| = n and $|B| = 2^n$ such that all vertices of B have pairwise different neighbourhood in A. Also, let S_n be the split graph obtained from Q_n by creating a clique in A.

Lemma 5. The graph Q_n is an n-universal bipartite graph, i.e. it contains every bipartite graph with n vertices as an induced subgraph.

Proof. Let G be a bipartite graph with n vertices and with parts A and B of size n_1 and n_2 , respectively. By adding at most n_2 vertices to A, we can guarantee that all vertices of B have pairwise different neighbourhoods in A. Clearly, Q_n contains the extended graph and hence it also contains G as an induced subgraph.

Corollary 1. Every co-bipartite graph with at most n vertices is contained in \overline{Q}_n and every split graph with at most n vertices is contained both in S_n and in \overline{S}_n .

Lemma 6. If a set A shatters a set B with $|B| = 2^n$, then B shatters a subset A^* of A with $|A^*| = n$.

Proof. Without loss of generality we assume that B is the set of all binary sequences of length n. Then every vertex $a \in A$ defines a Boolean function of n variables (the neighbourhood of a consists of the binary sequences, where the function takes value 1). For each $i = 1, \ldots, n$, let us denote by a_i the Boolean function such that $a_i(x_1, \ldots, x_n) = 1$ if and only if $x_i = 1$. Let A' be an arbitrary subset of $A^* = \{a_1, \ldots, a_n\}$ and $\alpha = (\alpha_1, \ldots, \alpha_n)$ its characteristic vector, i.e. $\alpha_i = 1$ if and only if $a_i \in A'$. Clearly, $\alpha \in B$ and $N(\alpha) \cap A^* = A'$. Therefore, B shatters A^* .

Lemma 7. For every n, there exists a k = k(n) such that every graph G with vc(G) = k contains one of $Q_n, \overline{Q}_n, S_n, \overline{S}_n$ as an induced subgraph.

Proof. Define $k = R(2^{R(n)})$, where R is the Ramsey number. Since vc(G) = k, there are two subsets A and B of V(G) such that |A| = k and B shatters A. By definition of k, A must have a subset A' of size $2^{R(n)}$ which is a clique or an independent set. Clearly, B shatters A' and hence, by Lemma 6, A' shatters a subset B' of B of size R(n). Then B' must have a subset B'' of size n which is either a clique or an independent set. Now $G[A' \cup B'']$ is either bipartite or co-bipartite or split graph, |B''| = n and A' shatters B''. Therefore, $G[A' \cup B'']$ contains one of $Q_n, \overline{Q}_n, S_n, \overline{S}_n$ as an induced subgraph.

Theorem 5. The classes of bipartite, co-bipartite and split graphs are the only three minimal hereditary classes of graphs of unbounded VC-dimension.

Proof. Clearly these three classes have unbounded VC-dimension, since they contain $Q_n, \overline{Q}_n, S_n, \overline{S}_n$ with arbitrarily large values of n.

Now let X be a hereditary class containing none of these three classes. Therefore, there is a bipartite graph G_1 , a co-bipartite graph G_2 and a split graph G_3 which are forbidden for X. Denote by n the maximum number of vertices in these graphs.

Assume that VC-dimension is not bounded for graphs in X and let $G \in X$ be a graph with vc(G) = k, where k = k(n) is from Lemma 7. Then G contains one of $Q_n, \overline{Q}_n, S_n, \overline{S}_n$, say Q_n . Since Q_n is n-universal for bipartite graphs (Lemma 5), it contains G_1 as an induced subgraph, which is impossible because G_1 is forbidden for graphs in X. This contradiction shows that VC-dimension is bounded in the class X.

4 More Results and Discussion

In [4], it was shown that for every t, p, s, there exists a z = z(t, p, s) such that every graph with a (not necessarily induced) matching of size at least z contains

either an induced matching of size t or an induced complete bipartite graph with color classes of size p or a clique of size s. This result was used in [4] to develop fpt-algorithms for the maximum induced matching problem in special classes of graphs where the problem is W[1]-hard. Now we use this result to derive a Ramsey-type characterization of the matching number $\mu(G)$, i.e. the size of a maximum matching in G. It is well known that $\mu(G) \leq \tau(G) \leq 2\mu(G)$, where $\tau(G)$ is the vertex cover number, i.e. the size of a minimum vertex cover in G. Therefore, the same characterization is valid the vertex cover number.

Theorem 6. The class of complete graphs, the class of complete bipartite graphs and the class \mathcal{M} of graphs of vertex degree at most 1 are the only three minimal hereditary classes of graphs of unbounded matching number and unbounded vertex cover number.

One more result of the same nature was proved in [2]. It states that for every t, p, s, there exists a z = z(t, p, s) such that every graph with a (not necessarily induced) path of length at least z contains either an induced path of length t or an induced complete bipartite graph with color classes of size p or a clique of size s. This result was used in [2] to obtain fpt-algorithms in special classes of graphs for the k-Biclique problem, which is generally W[1]-hard [8]. Now we use this result to derive the following Ramsey-type characterization of the *path number*, i.e. the length of a longest path.

Theorem 7. The class of complete graphs, the class of complete bipartite graphs and the class of linear forests (i.e. graphs every connected component of which is a path) are the only three minimal hereditary classes of graphs of unbounded path number.

From the last two theorems and the remark in the introduction it follows that path number lies between matching number and biclique number in the hierarchy of graph parameters. Also, it is well known that graphs of bounded path number have bounded tree-width and that the complete graphs and complete bipartite graphs are minimal hereditary classes of unbounded tree-width. Therefore, treewidth lies between path number and biclique number. On the other hand, it is known that tree-width lies below clique-width, i.e. bounded tree-width implies bounded clique-width, which was shown in [3]. In the same paper it was also shown that bounded clique-width together with bounded biclique number imply bounded tree-width. Therefore, the family of classes of bounded tree-width is precisely the intersection of the family of classes of bounded clique-width and the family of classes of bounded biclique number.

The above discussion shows that a Ramsey-type characterization of treewidth can be derived from such a characterization for clique-width, if it exists. However, in the case of clique-width even the existence of minimal classes is not obvious. The first two such classes have been recently discovered in [11]. In spite of this progress, a complete Ramsey-type characterization of clique-width is not possible, because in the universe of hereditary classes there are areas, where minimal classes do not exist, for instance, graphs of bounded vertex degree. There are several ways to overcome this difficulty. One of them is to reduce the universe. For instance, by reducing the universe of hereditary classes to minorclosed classes of graphs, we conclude that the class of planar graphs is the unique minimal minor-closed class of graphs of unbounded tree-width [13] and hence it is the unique minimal minor-closed class of graphs of unbounded clique-width. One more way to overcome the difficulty of non-existence of minimal classes is to employ the notion of boundary classes, which is a relaxation of the notion of minimal classes (see e.g. [10]).

Finally, instead of characterizing graph parameters in terms of "forbidden" elements (i.e. in terms of "what is not allowed") one can characterize them in terms of "what is allowed", and the results of the present paper suggest a uniform way to such characterizations. To see this, let us observe that both neighbourhood diversity and VC-dimension describe how complex the neighbourhood of a set X of vertices can be outside of X. This complexity can be described by a hypergraph whose hyperedges correspond to the neighbourhoods of vertices in X. In this terminology, neighbourhood diversity marks the jump from finitely many to infinitely many (distinct) hyperedges. Similarly, VC-dimension marks the jump from infinitely many hyperedges to all possible hyperedges. Between these two extremes, lies a variety of other graph parameters, such as cliquewidth, and exploring their neighbourhood complexity (i.e. the structure of the corresponding hypergraphs) is a very challenging task.

Acknowledgment. This work was supported by the Russian Science Foundation Grant No. 17-11-01336.

References

- Alon, N., Brightwell, G., Kierstead, H., Kostochka, A., Winkler, P.: Dominating sets in k-majority tournaments. J. Combin. Theory Ser. B 96, 374–387 (2006)
- Atminas, A., Lozin, V.V., Razgon, I.: Linear time algorithm for computing a small biclique in graphs without long induced paths. In: Fomin, F.V., Kaski, P. (eds.) SWAT 2012. LNCS, vol. 7357, pp. 142–152. Springer, Heidelberg (2012). https:// doi.org/10.1007/978-3-642-31155-0_13
- 3. Courcelle, B., Olariu, S.: Upper bounds to the clique width of graphs. Discrete Appl. Math. **101**, 77–114 (2000)
- Dabrowski, K.K., Demange, M., Lozin, V.V.: New results on maximum induced matchings in bipartite graphs and beyond. Theor. Comput. Sci. 478, 33–40 (2013)
- Gargano, L., Rescigno, A.: Complexity of conflict-free colorings of graphs. Theor. Comput. Sci. 566, 39–49 (2015)
- Hammer, P.L., Kelmans, A.K.: On universal threshold graphs. Comb. Probab. Comput. 3(3), 327–344 (1994)
- Lampis, M.: Algorithmic meta-theorems for restrictions of treewidth. Algorithmica 64(1), 19–37 (2012)
- Lin, B.: The parameterized complexity of k-Biclique. In: Proceedings of the Twenty-Sixth Annual ACM-SIAM Symposium on Discrete Algorithms, pp. 605– 615 (2015)
- Lozin, V., Rudolf, G.: Minimal universal bipartite graphs. Ars Comb. 84, 345–356 (2007)

- Lozin, V.: Boundary classes of planar graphs. Comb. Probab. Comput. 17(2), 287– 295 (2008)
- 11. Lozin, V.: Minimal classes of graphs of unbounded clique-width. Ann. Comb. **15**(4), 707–722 (2011)
- Ramsey, F.P.: On a problem of formal logic. Proc. Lond. Math. Soc. 30, 264–286 (1930)
- Robertson, N., Seymour, P.D.: Graph minors. V. Excluding a planar graph. J. Comb. Theory Ser. B 41, 92–114 (1986)