A TREE VERSION OF KÖNIG'S THEOREM

RON AHARONI, ELI BERGER, AND RAN ZIV

ABSTRACT. König's theorem states that the covering number and the matching number of a bipartite graph are equal. We prove a generalization, in which the point in one fixed side of the graph of each edge is replaced by a subtree of a given tree. The proof uses a recent extension of Hall's theorem to families of hypergraphs, by the first author and P. Haxell [?]. As an application we prove a special case (that of chordal graphs) of a conjecture of B. Reed.

1. WIDTH AND MATCHING WIDTH OF FAMILIES OF TREES

A *matching* in a hypergraph is a set of disjoint edges. In [?] (see also [?]) there were introduced a few notions which proved to be fruitful in the study of matchings. The two main ones are those of *width* and *matching width*, which are presented below.

Let H and F be two hypergraphs on the same vertex set. A subset C of F is said to be *H-covering* (or "an *H-covering*") if every edge in H meets some edge from C (in other words, the union of C is a vertex cover for H). The *F-width* $w(H,F)$ of H is the minimal size of an H-covering set of edges from F. The *F-matching width* $mw(H,F)$ of H is the maximum, over all matchings M in H, of $w(M,F)$. We write $w(H)$ and $mw(H)$ for $w(H,H)$ and $mw(H,H)$, respectively, and call these parameters the *width* and *matching width* of H. These notions have counterparts in which the covering set is required to be a matching. The *independent F-width* $iw(H,F)$ of H is the minimal size of a covering matching in F. The *independent matching width* $imw(H,F)$ of H is the maximum, over all matchings M in H, of $iw(M,F)$. Again, we write $iw(H)$ for $iw(H,H)$ and $imw(H)$ for $imw(H,H)$, and call them the *independent width* and *independent F-matching width* of H, respectively. (The notion of independent matching width is used in [?], but is not given there a special name.)

In [?] the following theorem was proved:

Theorem 1.1. *If $\mathcal{A}$ is a family of hypergraphs such that $mw(\bigcup\mathcal{B}) \geq |\mathcal{B}|$ for every $\mathcal{B} \subseteq \mathcal{A}$, then there is a choice of disjoint edges, one from each hypergraph in $\mathcal{A}$.*

Clearly, $mw \leq w$ and $imw \leq iw$, and in general strict inequalities occur. However, as we shall prove below, for families of trees equality holds in both. In fact, both equalities are special cases of a more general result. For two hypergraphs H, F and a symmetric relation $\sim$ on F we denote by $w(H, F, \sim)$ the minimal size of a $\sim$-related subset of F which is an H-covering. By $mw(H, F, \sim)$ we denote the maximum, over all matchings M in H, of $w(M, F, \sim)$.

The research of the first author was supported by the M. & M.L Bank Mathematics Research Fund and by the fund for the promotion of research at the Technion.

Theorem 1.2. *Let $H \subseteq F$ be families of subtrees of a given tree T. Let also $\sim$ be a symmetric relation on F containing the disjointness relation (i.e, if x, y are disjoint members of F then $x \sim y$). Then*

$$mw(H, F, \sim) = w(H, F, \sim)$$

The equality $mw(H) = w(H)$ for a family H of trees is obtained upon taking $F = H$ and taking $\sim$ to be the total relation (i.e., $x \sim y$ for all $x, y \in H$). (This equality can be formulated as stating that in a chordal graph it is as hard to dominate all the graph as it is hard to dominate independent sets in the graph.) The equality $imw = iw$ is obtained upon taking $F = H$ and $\sim$ the disjointness relation.

The theorem will clearly follow from the following lemma:

Lemma 1.3. *Let H, F and $\sim$ be as above, and let c, d be two intersecting elements of H. Then one of $w(H \setminus \{c\}, F, \sim)$ or $w(H \setminus \{d\}, F, \sim)$ is equal to $w(H, F, \sim)$.*

(Once the lemma is proved, the theorem follows by deleting edges from H one by one, until a matching is obtained).

Proof of the Lemma Write $w = w(H, F, \sim)$. Suppose, for contradiction, that $w(H \setminus \{c\}, F, \sim) = w(H \setminus \{d\}, F, \sim) = w - 1$. Let C be a $\sim$-related covering of size $w - 1$ of $H \setminus \{d\}$ and D a $\sim$-related covering of size $w - 1$ of $H \setminus \{c\}$. Clearly, D does not meet c and C does not meet d. Let S_1 be the connected component of $T \setminus (V(c) \cap V(\bigcup C))$ containing d and let $S_2 = T \setminus V(S_1)$. (Here $T \setminus A$, where A is a subset of $V(T)$, denotes the forest obtained by deleting the vertices of A from T.)

Let $C_i = \{g \in C : g \subseteq S_i\}$ $(i = 1, 2)$ and $D_i = \{g \in D : g \subseteq S_i\}$ $(i = 1, 2)$. Let also $K = C_1 \cup D_2 \cup \{c\}$ and $L = C_2 \cup D_1$.

From the definitions of S_1 and S_2 there follows:
Observation: A subtree of T meeting S_1 and S_2 contains a vertex in $\bigcup\{V(t) : t \in C\} \cap V(c) \cap V(S_2)$.

If a tree $t \in C$ meets both S_1 and S_2, then by the definition of S_1 it contains also points in $S_1 \setminus V(c)$. But this necessitates meeting d, which as noted above is impossible. Thus no $t \in C$ meets both S_1 and S_2, which means that $C_1 \cup C_2 = C$. Since a tree in D cannot meet c, by the above observation it cannot meet both S_1 and S_2, and thus also $D_1 \cup D_2 = D$.

From this, and from the choice of C and D, it follows that a tree in $H \setminus \{d\}$ covered neither by C_1 nor by D_2 must meet both S_1 and S_2, By the observation, this implies meeting c. Thus all trees in $H \setminus \{d\}$ are met by K. Since d is met by c, this means that K is an H-covering. Since D does not meet c, the tree c is $\sim$-related to all members of D_2. Also, by the definitions of S_1 and C_1, no tree from C_1 meets c, and hence $c \sim t$ for every $t \in C_1$. Thus K is $\sim$-related.

Take a tree $t \in H$. If t is contained in S_1 then it is different from c, and since D covers $H \setminus \{c\}$, it is met by a tree from D_1. If t is contained in S_2 then it is met by some tree in C_2. If t meets both S_1 and S_2, then by the above observation it is met by some tree $s \in C$ at a point of $V(c) \cap V(S_2)$. Since s contains a point from S_2, it is in C_2. Thus L is H-covering. Since D_1 and C_2 are each $\sim$-related, and since every tree in D_1 is disjoint from every tree in C_2 (and hence $\sim$-related to it), it follows that L is $\sim$-related.

We have shown that both K and L are $\sim$-related H-coverings. But $|K|+|L| = 2w-1$, and hence either $|K| < w$ or $|L| < w$. This yields a contradiction to the definition of w. $\square$

2. A König-like theorem

A hypergraph H is called a *point-tree* hypergraph if it is obtained from a bipartite graph by replacing, in each edge, the point in one side of the graph (the same side for all edges) by a tree. More formally, H is point-tree if, for some set X and a tree T whose vertex set is disjoint from X, each edge $e \in H$ is of the form $\{x\} \cup V(t)$, where $x = x(e) \in X$ and $t = t(e)$ is a subtree of T. For such a hypergraph we denote by $\sigma(H)$ the number $w(H, F)$, where $F = F(H) = \{\{x(e)\} : e \in H\} \cup \{V(t(e)) : e \in H\}$. (That is, the edges of H are to be caught by vertices in X, or trees $t(e)$). Recall that the *matching number* $\nu(H)$ of a hypergraph H is the maximal size of a matching in H. We shall prove the following generalization of König's theorem:

Theorem 2.1. *$\sigma(H) \leq \nu(H)$ for every point-tree hypergraph H.*

König's theorem is the special case in which all trees $t(e)$ are singletons.

Similar objects to point-tree hypergraphs are discussed in the literature. Notably, there have been considered hypergraphs obtained from a bipartite graph by replacing each point in each edge by an interval (the points in each side of the graph being replaced by subintervals of a given fixed interval). In [?](see also [?]) it was proved that $\tau \leq 2\nu$ for such hypergraphs. It is interesting to note that the proof in [?] was topological (as was the simpler proof in [?]), and at base so is the proof here, since it relies on Theorem ??, whose proof is topological.

In fact, we shall need a stronger version of Theorem ??, its "deficiency version".

Definition 2.2. *The* deficiency *$def(\mathcal{A})$ of a family $\mathcal{A}$ of hypergraphs is the minimal natural number d such that $mw(\bigcup \mathcal{B}) \geq |\mathcal{B}| - d$ for every subfamily $\mathcal{B}$ of $\mathcal{A}$.*

The following strengthening of Theorem ?? follows from it by a standard argument (see [?]).

Theorem 2.3. *Every family $\mathcal{A}$ of hypergraphs has a subfamily $\mathcal{D}$ of size at most $def(\mathcal{A})$, such that $\mathcal{A} \setminus \mathcal{D}$ has a choice function of disjoint edges.*

Proof of Theorem ?? For each $x \in X$ let $K(x)$ be the hypergraph $\{V(t(e)) : x(e) = x\}$. Applying Theorem ?? to the family $\{K(x) : x \in X\}$ yields the existence of a subset Y of X such that, writing $K = \bigcup\{K(y) : y \in Y\}$, we have $mw(K) \leq |Y| - (|X| - \nu(H))$. By Lemma ?? $w(K) = mw(K)$. Let C be a K-covering set of edges in K of size $|Y|-(|X|-\nu(H))$. Then $C \cup X \setminus Y$ is an H-covering set consisting of elements of $F(H)$, of size $\nu(H)$. This proves the theorem. $\square$

3. The special case of families of intervals

In this section we consider a special case of the general setting discussed above. It is the case in which the big tree is an interval, and of course all the subtrees are intervals. It turns out that in this case the equalities $mw = w$ and $imw = iw$ can be strengthened. For example, the matchings which are "hard to catch" can be described explicitly, and can be found efficiently.

Here is some notation which we shall use. We say that a set of vertices in a graph G is *k-independent* if the distance between any two of its members is larger than k. (Thus 1-independence is just the usual independence.) By $\gamma_k(G)$ we denote the maximal size of a k-independent set in G. A set of edges in a hypergraph H is *k-remote* if it is k-independent in the line graph $L(H)$. We write $\zeta_k(H)$ for $\gamma_k(L(H))$. Clearly, for all hypergraphs we have

$$\zeta_2 \leq mw \leq w \tag{1}$$

Here we shall prove:

Lemma 3.1. *For a family F of intervals on the real line equality holds throughout (??), namely $\zeta_2(F) = mw(F) = w(F)$.*

Given a family of intervals F, we denote by $\ell(F)$ the leftmost interval in F, namely the one with minimal right endpoint. For a real number x we denote by $F(> x)$ the set of those intervals in F which lie entirely within (x, ∞).

Proof of Lemma ?? By (??) the lemma will be proved if we can exhibit a 2-remote set R and a covering set C of the same size. We construct such sets by an inductive process.

Let $r_1 = \ell(F)$, and $c_1 = [x_1, y_1]$ an element of F meeting r_1 whose right endpoint is largest. Let now $r_2 = \ell(F(> y_1))$, and $c_2 = [x_2, y_2]$ the interval in F meeting r_2 whose right endpoint is maximal. (Note that c_2 may intersect c_1.) In general, having defined $c_i = [x_i, y_i]$ $(i > 0)$, we let $r_{i+1} = \ell(F(> y_i))$ and choose c_{i+1} as the interval in F meeting r_{i+1} and extending furthest to the right. The process terminates when, at some stage t, there is no interval lying entirely to the right of c_t. Then, clearly, $R = \{r_1, \ldots, r_t\}$ is remote, while $C = \{c_1, \ldots, c_t\}$ is covering. $\square$

A similar equality can be proved for ζ_k for all k. For any graph G we denote by $\rho_k(G)$ the minimal number of vertex sets of diameter at most k, whose union is $V(G)$. Obviously, $\rho_k(G) \geq \gamma_k(G)$ for all k. One corollary of Lemma ?? is that in an interval graph $\rho_2(G) = \gamma_2(G)$ (in fact, it implies something even stronger, namely that an interval graph G can be partitioned into $\gamma_2(G)$ sets of radius 1.) This equality is true for all k, that is, in an interval graph $\rho_k(G) = \gamma_k(G)$ for all k. We shall prove this for a more general class of graphs, that of *incomparability graphs*. A graph is called an *incomparability graph* if there exists a partial order on its vertex set, such that two vertices in the graph are connected if and only if they are incomparable. As is well known, an interval graph is an incomparability graph, the order on the intervals being that of "lying completely to the right of".

Proposition 3.2. *In an incomparability graph $\rho_k(G) = \gamma_k(G)$ for all k.*

Given a graph G and a natural number k, we denote by $G^{\leq k}$ the graph with the same vertex set as G, in which two vertices are connected if and only if their distance in G does not exceed k. Obviously, $\gamma_k(G) = \gamma(G^{\leq k})$, and $\rho_k(G) = \rho(G^{\leq k})$. Thus, if $G^{\leq k}$ is perfect, then G satisfies Proposition ??. Since incomparability graphs are known to be perfect, Proposition ?? will follow from:

Lemma 3.3. *If G is an incomparability graph then so is $G^{\leq k}$.*

Proof Let "$\succ$" be a partial order on $V(G)$ such that x and y are connected in G if and only if they are $\succ$-incomparable. Let "$\sqsupset$" be the relation defined by $x \sqsupset y$ if and only if $x \succ y$ and $(x, y) \notin E(G^{\leq k})$. Clearly, it suffices to show that "$\sqsupset$" is a partial order.

Assume that it is not. Then there exist $x, y, z \in V(G)$ such that $x \sqsupset y$, $y \sqsupset z$, while $x \not\sqsupset z$. There exist then vertices $x = u_0, u_1, u_2, \ldots, u_t = z$ $(t \leq k)$ such that u_i is $\succ$-incomparable to u_{i+1} for all $0 \leq i < t$. Since $x \sqsupset y$, all u_i must be $\succ$-comparable to y (or else the sequence $u_0, \ldots, u_i, y$ would show that $(x, y) \in E(G^{\leq k})$). Since $u_0 \succ y$ and $y \succ u_t$, it follows that there exists $i < t$ such that $u_i \succ y$ and $y \succ u_{i+1}$. But then, by the transitivity of the relation $\succ$, we have $u_i \succ u_{i+1}$, a contradiction. $\square$

Our next observation on the interval case is that the equality $imw = iw$ can be given in this case a constructive (and algorithmically efficient) proof.

Theorem 3.4. *$imw(F) = iw(F)$ for every hypergraph F of intervals. Moreover, there is a polynomial time algorithm producing a matching M such that $iw(M, F) = iw(F)$.*

Proof Write $k = iw(F)$. We shall prove the theorem by constructing a matching M in F satisfying:

(P) Every matching of size $k - 1$ in F misses an interval from M.

Note, first, that we can assume that no two endpoints of intervals in F coincide: if there is such coincidence, we can enlarge some intervals, without changing the intersection pattern of the family. Henceforth we shall make this assumption.

For a matching E in F we write $E = (e_1, \ldots, e_t)$ if e_i are the edges of E ordered from left to right. Let E be such a matching, and write $e_j = [a_j, b_j]$, $j \leq t$. For each $j \leq t$ we denote by $d_j(E)$ the interval $\ell(F(> b_{j-1}))$ (where b_0 is defined as $-\infty$, meaning that $d_1(E)$ is always $\ell(F)$). If $F(> b_t)$ is non-empty, we write also $d_{t+1}(E) = \ell(F(> b_t))$. Note that $d_j(E)$ meets e_j if and only if there is no interval from F lying between e_{j-1} and e_j; or, in the case $j = 1$, if and only if there is no interval in F lying to the left of e_1.

We denote by $m(E)$ the maximal index $m \leq t + 1$ such that there is no interval from F lying between e_{j-1} and e_j for all $j < m$ (if there is an interval from F to the left of e_1, we write $m(E) = 1$). If $m(E) = t + 1$, we say that E is *dense*. Note that if E misses some edge from F then it misses $d_{m(E)}(E)$.

Let L_i be the set of all right endpoints of intervals e_i in some dense matching $E = (e_1, \ldots, e_t)$. An interval $[a, b]$ is called *L_i-free* if $[a, b] \cap L_i \subseteq \{b\}$. Let M be the set of all intervals in F which are equal to $d_i(E)$ for some dense matching E, and are L_j-free for all $j \leq i$.
The proof will be complete if we show that M is a matching, and that it satisfies property (P).

Let us first show that M is a matching. Suppose, for contradiction, that $d = d_p(E) = [a, b]$ and $d' = d_q(E') = [a', b']$ are two intersecting members of M, where $E = (e_1, \ldots, e_{p-1})$ and $E' = (e'_1, \ldots, e'_{q-1})$ are two dense matchings. By assumption $b \neq b'$, so we may assume (say) that $b < b'$. Note that $b \in L_p$, since the matching $(e_1, \ldots e_{p-1}, d)$ is dense. Since $d' \in M$, this implies that $q < p$. Let y be the right endpoint of e'_{q-1}. Since $d \in M$ and $y \in L_{p-1}$, it follows that $y < a$. But then the interval $d = [a, b]$ witnesses the fact that $d' \neq \ell(F(> y))$, contradicting the definition of d'.

Next we show that M satisfies property (P). Suppose that there exists a matching $Z = (z_j : j \leq t)$ in F meeting all intervals in M, where $t < k$. Among all such matchings Z, choose one in which the right endpoint of $z_{m(Z)}$ is maximal. Write

$m = m(Z)$ for this Z. Since $iw(F) = k$, the matching Z misses an interval from F. This means that Z misses $d = d_m(Z)$. If $d \in M$, we are done. If not, then d contains a point from L_j for some $j \leq m$. That is, it is the right endpoint of an interval w_j in some dense matching $W = (w_1, \ldots, w_m)$. But then the matching $(w_1, \ldots, w_j, z_{m+1}, \ldots, z_t)$, like Z, does not miss any edge from M. This contradicts the maximality property of Z.

To see that M above can be constructed in polynomial time, it suffices to note that the sets D_j of those intervals which appear as the j-th element in some dense matching can be constructed inductively in plynomial time, and that M is defined by these sets. $\square$

It is well known (see e.g. [?]) that any chordal graph has a "tree representation", that is, it is the intersection graph of a family of subtrees of some tree. Hence it is easy to derive from Theorem **??** the following corollary:

Corollary 3.5. *Let G be a chordal graph, and let $A_1, \ldots, A_m$ be sets of vertices in G. If for every subset I of $[m]$ there obtains $dom(G[\bigcup_{i \in I} A_i]) \geq |I|$, then there exists a choice of distinct elements $a_1 \in A_1, a_2 \in A_2, \ldots a_m \in A_m$, such that $\{a_i : i \leq m\}$ is independent.*

(Here by $dom(H)$ we mean the domination number of the graph H, that is the minimal number of vertices needed to dominate H. The derivation of Corollary **??** from Theorem **??** is the same as that of Hall's theorem from König's theorem.)

4. An application to a conjecture of B. Reed

B. Reed [?] made the following conjecture:

Conjecture 4.1 (Reed)**.** *Let G be a graph, and let L_v be a list of colors of size $k+1$ given at each vertex v of G. Suppose that for every color c the subgraph of G induced by the vertices v for which $c \in L_v$ has degree at most k. Then there is a legal list-coloring of the graph from the given lists, i.e a choice of a color $c_v \in L_v$ for every vertex v, such that $c_u \neq c_v$ whenever u and v are adjacent.*

Let us first strengthen this conjecture. With the graph G and lists L_v as in the conjecture, let us say that a set of pairs $\{(u_i, c_i) : i \in I\}$ *dominates* a set U of vertices of G if $u_i \in U$ and $c_i \in L_{u_i}$ for every $i \in I$, and for every vertex $v \in U$ and color $d \in L_v$, for some $i \in I$ the vertex u_i is adjacent or equal to v, and $c_i = d$.

Reed's conjecture would clearly follow from the following one:

Conjecture 4.2. *Let G and the lists L_v be as above. If no set U of vertices of G is dominated by fewer than $|U|$ pairs, then there is a legal list-coloring of the graph from the given lists.*

In [?] it was shown how hypergraph matching results can be applied to Reed's conjecture. We shall use the same idea here to prove Conjecture **??** (and hence also Conjecture **??**) for chordal graphs.

Let G be a chordal graph, and L_v be lists of colors at its vertices, satisfying the condition of the conjecture. Let H be the graph on all pairs (u, c) such that $c \in L_u$, and in which (u, c) is adjacent to (v, d) if and only if $c = d$ and u, v are adjacent in G. Then H is the disjoint union of induced subgraphs of G, and thus it is chordal.

For each vertex v of G let $A_v = \{(v, c) : c \in L_v\}$. By the condition of the theorem, for every subset U of $V(G)$ we have $dom(\bigcup_{u \in U} A_u) \geq |U|$. By Corollary

?? it follows that there exists a choice function f of elements from A_v, whose range is independent in H. Coloring each vertex v by the color c such that $f(v) = (v, c)$ yields a proper list coloring, as desired. $\square$

References

[1] R. Aharoni, Ryser's conjecture for tripartite hypergraphs, *Combinatorica*, to appear

[2] R. Aharoni and P. Haxell, Hall's theorem for hypergraphs, *Jour. Graph Th.*, to appear

[3] F. Gavril, The intersection graphs of subtrees in trees are exactly the chordal graphs, *J. Combin. Th. Ser.B* 16(1974),47–56.

[4] P. Haxell, A note on vertex list colouring, *submitted for publication*

[5] T. Kaiser, Transversals of d-intervals, *Discrete. Comput. Geom.* 18 (1997), 195–203.

[6] R. Meshulam, The clique complex and hypergraph matching, *Combinatorica*, to appear

[7] B. Reed, The list colouring constants, *Journal of Graph Theory* 31(1999), 149–153.

[8] G. Tardos, Transversals of 2-intervals, a topological approach, *Combinatorica* 15 (1995), 123–134.

Department of Mathematics, Technion, Haifa, Israel 32000
E-mail address, Ron Aharoni: `ra@tx.technion.ac.il`

Department of Mathematics, Technion, Haifa, Israel 32000
E-mail address, Eli Berger: `eberger@princeton.edu`

Department of Computer Science, Tel-Hai college, Upper Galilee, Israel 12210
E-mail address, Ran Ziv: `ranziv@telhai.ac.il`