The Expansion Theorem for Median Graphs

Henry Martyn Mulder

To Cor Baayen, at the occasion of his retirement

1. Introduction

This paper deals with the adventures of the expansion theorem for median graphs [Mu 78, Mu 80b]. It was the first theorem I ever proved (on a walk during the Fifth Hungarian Combinatorial Conference in Keszthély, Hungary, I was 'struck' by the idea ultimately leading to this theorem), and it was the starting point for my Ph. D. thesis written under the inspiring guidance of Cor Baayen. Median graphs existed already in the literature [Av 61, Ne 71], but they were independently introduced by LEX SCHIJVER and me [MS 79] in the context of a problem in finite topology posed by JAN VAN MILL [vM 77] early 1976. All three of us were at that time Ph. D. students of Cor Baayen.

Loosely speaking the idea of expansion is the following. Let G be covered by a number of subgraphs, which two by two intersect in the same subgraph G_0 of G. Now we take disjoint copies of the covering subgraphs and join the respective copies of G_0 in these subgraphs by new edges.

By imposing conditions on the covering subgraphs and on how to insert the new edges we get specific instances of expansion. Some types of expansion may not be sensible to study, but others seem to be quite promising in producing interesting problems and results. In [Mu 90] a 'masterplan' was formulated for studying various expansion problems.

To show how fruitful this approach can be, we discuss a number of results on median graphs. These all have elegant and straightforward proofs using a specific instance of expansion, by which median graphs can be characterized. A median graph is a graph such that, for any triple of vertices u,v,w, there exists a unique vertex minimizing the sum of the distances to u, v and w.

2. Median graphs and expansion

In this section, we will give some results and introduce terminology found in [Mu 78, Mu 80b, Mu 90, MMR 94]. All graphs considered in this paper will be finite, and we use the standard notation $G=(V,E)$ to denote a graph with vertex set V and edge set E. We will often simply write only G and leave

V and E understood. Also, we will not distinguish between a subset W of V and the subgraph induced by W. In a connected graph, the ***distance*** $d(x,y)$ between two vertices x and y is the length of a shortest x,y-path, or an ***x,y-geodesic***. The star of our show is the ***median graph*** G: a connected graph such that for every three vertices x,y,z, of G, there is a unique vertex w on a geodesic between each pair of x,y,z. This vertex w is called the ***median*** of the triple x,y,z. The ***interval*** between the vertices x and y is the set $I(x,y)$ of all vertices on x,y-geodesics, i.e.,

$$I(x,y) = \{w \in V : d(x,w) + d(w,y) = d(x,y)\}.$$

The ***interval function*** I of a graph G was extensively studied in [Mu 80b]. It is an easy observation that a graph G is a median graph if and only if $|I(x,y) \cap I(y,z)| = 1$ for all vertices x,y,z of G. Median graphs were first studied in 1961 by AVANN [Av 61], and independently introduced by NEBESKY [Ne 71] and MULDER and SCHRIJVER [MS 79]. Trees are the simplest examples of median graphs. Another prime example is the n-cube Q_n. Recall that Q_n has $\{0,1\}^n$ as vertex set, and two vertices are adjacent whenever they differ in exactly one place. For three vertices $x = x_1x_2...x_n, y = y_1y_2...y_n, z = z_1z_2...z_n$ of Q_n the median $w = w_1w_2...w_n$ of x,y,z is determined by the majority rule: $w_i = \delta$ if δ occurs at least twice among x_i, y_i, z_i, for $i = 1,...,n$. Other examples of median graphs are the grids and the covering graphs of distributive lattices. It is also an easy observation that median graphs are bipartite, for if $x_0...x_kx_{k+1}...x_{2k}x_0$ is a shortest cycle of odd length, then x_0, x_k, x_{k+1} would have x_k and x_{k+1} as two distinct medians. The smallest bipartite graph that is not a median graph is $K_{2,3}$: the profile consisting of three independent vertices has two medians.

A set W of vertices of a graph G is ***convex*** if $I(x,y) \subseteq W$ for every $x,y, \in W$, and a ***convex subgraph*** of G is a subgraph induced by a convex set of vertices of G. Clearly, a convex subgraph of a connected graph is also connected. Moreover, the intersection of convex sets (subgraphs) is convex. The ***convex hull*** $Con(U)$ of a set of vertices U is the intersection of all the convex sets containing U. It was proved in [Mu 80b] that intervals in median graphs are convex, so that $Con(\{x,y\}) = I(x,y)$. Also, in median graphs, convex sets can be viewed in another useful way through the notion of a gate. For $W \subseteq V$ and $x \in V$, the vertex $z \in W$ is a ***gate*** for x in W if $z \in I(x,w)$ for all $w \in W$. Note that a vertex x has at most one gate in any set W, and if x has a gate z in W, then z is the unique nearest vertex to x in W. The set W is ***gated*** if every vertex has a gate in W and a ***gated subgraph*** is a subgraph induced by a gated set. It is not difficult to see that in any graph, a gated set is convex and that in a median graph a set is gated if and only if it is convex. (This last fact follows immediately from results in [Mu 80b].)

Recall that for two graphs $G_1=(V_1,E_1)$ and $G_2=(V_2,E_2)$, the ***union*** $G_1 \cup G_2$ is the graph with vertex set $V_1 \cup V_2$ and edge set $E_1 \cup E_2$, and the ***intersection*** $G_1 \cap G_2$ is the graph with vertex set $V_1 \cap V_2$ and edge set $E_1 \cap E_2$. We write $G_1 \cap G_2 = \emptyset$ $(\neq \emptyset)$ when $V_1 \cap V_2 = \emptyset (\neq O)$. A *proper cover* of G consists of two convex subgraphs G_1 and G_2 of G such that $G = G_1 \cup G_2$,$G_1 \cap G_2 \neq \emptyset$. Every graph G admits the ***trivial proper cover*** G_1 ,G_2 with $G_1 = G_2 = G$. On the other hand a cycle does not have a proper cover with two proper subgraphs.

We are now able to give the definition of the operation which will help yield a characterization of median graphs. Let $G'=(V',E')$ be properly covered by the convex subgraphs $G_1'=(V_1',E_1')$ and $G_2'=(V_2',E_2')$ and set $G_0'=G_1' \cap G_2'$. For $i=1,2$, let G_1 be an isomorphic copy of G_1', and let λ_i be an isomorphism from G_i' onto G_1. We set $G_{0i}=\lambda_i[G_0']$ and $\lambda_i(u')=u_i$, for u' in G_0'. The ***expansion*** of G' with respect to the proper cover G_1', G_2' is the graph G obtained from the disjoint union of G_1 and G_2 by inserting an edge between u_1 in G_{10} and u_2 in G_{20}, for each u' in G_0'. Denote the set of edges between G_{10} and G_{20} by F_{12}. This is illustrated in Figure 1. We say that λ_i ***lifts*** G_i' up to G_i. For any subgraph H' of G' we abuse the notation and write $\lambda_i[H']$ for $\lambda_i[H' \cap G_i']$. So λ_i lifts the part of H' lying in G_i' up to G_i.

This type of expansion was called a "convex expansion" in [Mu 78], [Mu 80b], and a "convex Cartesian expansion" in [Mu90] for a more general setting. We are now able to state the following fundamental result on median graphs first proved in [Mu 78] and [Mu 80b]. This result is the basis of a recent $O(|V|^2 \log|V|)$ algorithm for recognizing median graphs found in [JS].

Theorem 1. A graph G is a median graph if and only if G can be obtained by successive expansions from the one vertex graph K_1.

Using this theorem, trees can be obtained from K_1 by restrictng the expansions to those of the following type: G_1 is always the whole graph G and G_2 is a single vertex. Expansio with respect to such a cover amounts to adding a new vertex adjacent to the one in G_2. The n–cubes can be obtained from K_1 by using only trivial proper covers. Note that $K_{2.3}$ can not be obtained from a smaller graph by expansion with respect to a proper cover.

In order to make full use of Theorem 1 and to develop additional techniques, we give a very brief sketch of the proof. Along the way we introduce some extra terminology.

The basic ideas used for the proof of Theorem 1 are the following. Take an arbitrary edge v_1v_2 in a median graph G. Let G_1 be the subgraph of G induced by all vertices nearer to v_1 than to v_2, and let G_2 be the subgraph induced by all vertices nearer to v_2 than v_1. Since G is bipartite, it follows that G_1, G_2 partition G. We call such a partition a *split*. Let F_{12} be the set of edges between G_1 and G_2, and let G_{i0} be the subgraph induced by the endvertices in G_i of the edges in F_{12}, for $i=1,2$. Then one proceeds to prove the following facts (not necessarily in this order):

(i) F_{12} is a matching as well asa cutset (minimal disconnecting edge–set).

(ii) The subgraphs G_1,G_2,G_{10},G_{20} are convex subgraphs of G.

(iii) The obvious mapping of G_{10} onto G_{20} defined by the edges in $F_{12}(u_1 \to u_2$, for any edge u_1u_2 in F_{12} with u_i in G_{i0}, for $i=1,2)$ is an isomorphism.

(iv) For every edge u_1u_2 of F_{12} with u_i in G_{0i} $(i=1,2)$, the subgraph G_1 consists of all the vertices of G nearer to u_1 than to u_2, so that u_1 is the gate in G_1 for u_2. A similar statement holds for G_2.

Now the *contraction* G' of G with respect to the split G_1,G_2 is obtained from G by contracting the edges of F_{12}. To illustrate this in Figure 1, move from right to left. Clearly expansion and contraction are inverse operations. The *contraction map* κ, of G onto G', associated with F_{12} is thus defined by $\kappa|_{G_i}=\lambda_i^{-1}$, for $i=1,2$. Finally one shows that G' is a median graph and so, by induction on the number of vertices, Theorem 1 is proved.

We present another feature of median graphs that helps in getting the right mental picture of how to operate with them in the rest of the paper. A *cutset coloring* of a connected graph is a proper colouring of the edges (adjacent edges have different colours) such that each colour class is a cutset (a minimal disconnecting edge set). Of course, most graphs will not have a cutset colouring, whereas even cycles of length at least six have more than one. if we want to cutset colour the edges of a graph, then in an induced 4–cycle $wxyzw$, opposite edges most have the same colour. So, w, z are on one side and x, y on the other side of the cutset colour of wx, and thus yz gets the same colour as wx. We call this the *4–cycle property* of cutset colourings. It follows from (i), (ii) and (iii) that in any cutset colouring of the median graph G, the set F_{12} must be a colour class. Using induction on the number of colours gives the next corollary [Mu 78, Mu 80b].

Corollary 2. A median graph is uniquely cutset colourable up to the labeling of the colours.

For a split G_1,G_2, we call the set F_{12} a *colour*, and G_1 and G_2 the *colourhalves* of F_{12}. Thus it follows that any colour in the cutset colouring of the median graph G defines a split into two convex colourhalves, as in the case of F_{12} with all the properties listed above. Hence the 4-cycle property, one can determine the colour class of an arbitrary edge xy. This colour class splits G into the convex subgraph of all vertices nearer to x than to y and the convex subgraph of all vertices nearer to y than to x, etc. There is yet another important feature of median graphs that we need in the sequel, and which follows from (the proof of) Theorem 1 [Mu 80b]. If we consider any two colours in the cutset colouring of the median graph G , and we contract them in any order, then we get the same median graph G''. Hence we can apply the corresponding expansions to obtain G from G'' in any order. This means that in obtaining G from a median graph H by a succession of expansions, we can apply these expansions in any order. This is easily seen in the case for trees: every expansion corresponds to an edge in the tree, and it does not matter in what order we introduce the edges in forming the tree.

The basic technique that will be used in proofs found in the next section is as follows: One or more contractions on the median graph G are performed to obtain a smaller median graph G', on which we apply the appropriate induction hypothesis. Then we perform the corresponding expansions in reverse order on G' so that we regain G. During this process, a vertex x of G is contracted to a unique vertex x' in G'. When we recover G from G' by expansions, then x' is lifted up in each expansion to the appropriate colourhalf until we regain x. The sequences of vertices and expansions that we obtain in this way from x' up to x is called the *history* of x (with respect to the expansions involved). Similarly, if $\pi = (x_1,\dots,x_k)$ is a sequence of vertices of G, a *profile* for short, then π is contracted to a profile $\pi' = (x_1',\dots,x_k')$ on G', where x_i' is the contraction of x_i, for $i = 1,\dots,k$. We thus define the *history* of π in the obvious way. If x' is a vertex of G' and we lift x' up to a vertex x in an expansion of G', then we call x a *descendant* of x'. Hence if we know which lifts are applied on x' in the expansions to regain G from G', then we know the history of all the descendants of x'.

Having now introduced the basic techniques and results on median graphs, we will use them frequently without specific mention in the sequel.

In a sense median graphs are the appropriate common generalization of trees and hypercubes. This as well as many other results on median graphs suggest the following 'meta'conjecture.

Metaconjecture. Any 'reasonable' property shared by trees and hypercubes is shared by all median graphs.

In the rest of the paper, we use the standard notation developed above: G is a median graph with split G_1, G_2 with colour $F_{12,}$ contraction G' etcetera.

3. Median sets

Let G be a connected graph. A *profile* on G is W is a vertex sequence $\pi=(v_1,v_2,\dots,v_k)$ in G. Note that multiple occurrences in π are allowed. The *length* k of the profile is denoted by $|\pi|$. A profile is *even* or *odd* depending on whether k is even or odd. The (*simultaneous*) *distance* $D(u,\pi)$ of a vertex u to π is defined by

$$D(u,\pi) \;=\; \sum_{i=1}^{k} d(u,v_i).$$

A *median* of π is a vertex x minimizing the distance $D(x,\pi)$, and the *median set* $M(\pi)$ of π consists of the medians of π. Since G is assumed to be connected, a median set is always non-empty. The median set of two vertices u,v is the interval $I(u,v)$. In general not much is known about the structure of median sets, but not so for median graphs. Clearly here every triple of vertices has a unique median. For longer profiles the situation is equally plain. After one has made the effort to develop the expansion technique, on can sit down in the armchair and let the expansions do the work. In [MMR 94] the expansion technique is exploited in its full richness to study median sets in median graphs. We present here the main results and prove one Lemma to give an idea how one could proceed to prove the theorems.

If π is a profile in a median graph G with split G_1, G_2, then let π_i be the subprofile of π consisting of all elements of π in G_i. For each subset W of V, we set $W'=\kappa[W]$ and $x'=\kappa(x)$. Note that if for some u' in G_0', both u_1 and u_2 are elements of W, them u' is in W' and $|W'|<|W|$. If π is a profile on G, then we have $\pi_i'=\kappa(\pi_i)$ and $\pi_i=\lambda_i(\pi_i')$, where κ and λ_i are applied componentwise.

Lemma 3. With the above notation, if π is a profile in the median graph G with $|\pi_1|>|\pi_2|$, then $M(\pi')$ is contained in G_1', and $M(\pi)$ is contained in G_1, and $M(\pi')=M(\pi)'$, and $|M(\pi)|=|M(\pi')|$.

Proof. Let w' be a vertex in $G_2' - G_0'$ and let x' be the gate of w' in G_1'. Then we have

$$D(w',\pi_1') = D(x',\pi_1') + |\pi_1'|d(x',w').$$

The triangle inequality for d yields

$$D(w',\pi_2') \geq D(x',\pi_2') - |\pi_2'|d(x',w').$$

Hence we have

$$\begin{aligned} D(w',\pi') &= D(w',\pi_1') + D(w',\pi_2') \\ &\geq D(x',\pi') + d(x',w')(|\pi_1'| - |\pi_2'|) \\ &> D(x',\pi'). \end{aligned}$$

So $M(\pi')$ lies in G_1'.
Now choose a vertex w in G_2 and a vertex v in G_1 with v' in $M(\pi')$. Then we have

$$D(v,\pi) = D(v',\pi') + |\pi_2|,$$

$$D(w,\pi) = D(w',\pi') + |\pi_1|,$$

whence $D(w,\pi) > D(v,\pi)$. So $M(\pi)$ lies in G_1. Finally, for each vertex v in G, we have

$$D(v,\pi) = D(v',\pi') + |\pi_2|,$$

so that $M(\pi)' = M(\pi')$. Since $M(\pi)$ lies in G_1, it follows that $|M(\pi)| = |M(\pi')|$. □

Using this Lemma, we can relate $M(\pi)$ to the median set of π' in G'.

Theorem 4. If π is a profile in a median graph G with $|\pi_1| > |\pi_2|$, then $M(\pi) = \lambda_1[M(\pi')]$. Furthermore, if π is odd, then $|M(\pi)| = 1$. If π is even, then $M(\pi)$ is an interval, and if $|\pi_1| = |\pi_2|$, then $M(\pi) = \lambda_1[M(\pi')] \cup \lambda_2[M(\pi')]$

Using expansions, we can also relate $M(\pi)$ to the median sets of its vertex-deleted subprofiles.

Theorem 5. Let $\pi = (v_1, v_2, \ldots, v_k)$ a profile in a median graph G with $k > 1$. If π is odd, then $M(\pi) = \bigcap_i M(\pi - v_i)$, and if π is even, then $M(\pi) = Con(\bigcap_i M(\pi - v_i))$.

For proofs the reader is referred to [MMR 94].

4. Dynamic search

In [CGS 87] and [CGS 89] CHUNG, GRAHAM and SAKS considered the following intriguing problem and proved some important results.

Let $G = (V,E)$ be a connected graph, where on each vertex some piece of information is located. A retriever is located at some vertex u of G, his *position*. A *quest* for a piece of information comes in the form of a quest for the vertex where this information is located. The retriever has two options:

(i) to retrieve from u the information at v, which costs $d(u,v)$;
(ii) to move from u to some vertex v, which also costs $d(u,v)$.

If the retriever is at an *initial position* p_0, then his goal is, given a sequence of quests $Q = q_1, q_2, \ldots, q_n$ to find a sequence of positions $P = p_0, p_1, \ldots, p_n$ such that the following distance sum is minimized:

$$(*) \qquad \sum_{i=1}^{n} d(p_{i-1}, p_i) + d(p_i, q_i).$$

We can read this sum as follows: being at p_{i-1}, the retriever first moves to p_i and then retrieves q_i, for $i = 1, \ldots, n$.

With each *quest sequence* Q and each *position sequence* P we can associate a *caterpillar* $R(P,Q)$ consisting of P, Q and a p_{i-1},p_i–geodesic and a p_i,q_i–geodesic, for $i = 1, \ldots, n$. Note that in $R(P,Q)$ we may have multiple occurrences of vertices as well as edges. The p_{i-1},p_i–geodesics with $i = 1, \ldots, n$ form the *spine* of the caterpillar, the p_i,q_i–geodesics are the *legs* (note that mathematics is capable of creating new biological species). The *length* $\ell(P,Q)$ of the caterpillar $R(P,Q)$ is the sum of the lengths of all geodesics involved in constructing the caterpillar, and thus $\ell(P,Q)$ equals sum (*) above. In these terms, given a quest sequence Q and initial position p_0 we want to find a *shortest* caterpillar $R(P,Q)$.

If the retriever being at the initial position knows all the quests in quest sequence Q, then he can always find a shortest caterpillar $R(P,Q)$

minimizing his total costs. How to find P is another story. But if he has only partial knowledge at some position p_{i-1}, he can only optimize p_i with respect to, say, the next k quests $q_i, q_{i+1}, \ldots, q_{i+k-1}$. When finally all quests have come in and he has completed his caterpillar $R(P,Q)$, then it is generally not the shortest possible caterpillar.

If at any position p_{i-1} we have only foreknowledge of the next two quests q_i and q_{i+1}, then the best thing we can do is choosing a median point of p_{i-1}, q_i, q_{i+1} as our next position p_i. This is the *median strategy*. In [CGS 87] the problem was posed and settled on which graphs the median strategy, with always foreknowledge of the next two quests at each position, will produce a shortest caterpillar for each initial position p_0 and each quest sequence Q, cf. [Wr 87].

Theorem 6. Let G be a connected graph. The median strategy with foreknowledge of the next two quests at each position produces a shortest caterpillar for each initial position p_0 and each quest sequence Q if and only if G is a median graph.

If the median strategy is optimal, then CHUNG, GRAHAM and SAKS proceed in the following way. Assume that there are vertices u,v,w having two distinct median points. Choose such a triple with $d(u,v)+d(v,w)+d(w,u)$ as small as possible. Now, with initial position u, by choosing quest sequences of length at most 6 of the type u,u,v,w,q,q and varying q, a contradiction can be derived. For full details of this proof the reader is referred to [CGS 87]. To prove the converse they make use of BANDELT'S theorem [Ba 84] that the median graphs are precisely the retracts of hypercubes (see the next subsection). Here we give an alternative proof for the 'if part' using our expansion approach.

Proof of the 'if part' of Theorem 6. We use induction on the number of expansions, so let $F, G_1, G_2, G', \pi', W'$ etcetera be as above. Let P be the position sequence obtained via the median strategy with respect to initial position p_0 and quest sequence Q. Note that, because of unicity of medians, P is uniquely determined.

Assume that there is a position sequence T with $\ell(T,Q)<\ell(P,Q)$. Note that P' is the position sequence obtained via the median strategy in G' with respect to p_0' and Q'. By induction hypothesis, we know that $\ell(T',Q')\geq\ell(P',Q')$. Note that, for any caterpillar $R(S,Q)$ in G, it follows from the expansion procedure that

$$\ell(S,Q) \;=\; \ell(S',Q')+\alpha(S,Q),$$

where $\alpha(S,Q)$ is the number of edges from F lying on $R(S,Q)$.

Without loss of generality we may assume that p_0 lies in G_1. Put $q_0=p_0$. The spine of $R(P,Q)$ starts in G_1. Beginning in p_0 we walk along the spine of $R(P,Q)$ and check where the caterpillar crosses the cut F:

- if p_{i-1}, q_{i-1}, q_{i+1} lie in G_1 and q_i lies in G_2, then the crossing is in the p_i,q_i–leg and the spine remains in G_1,
- if p_{i-1}, q_{i-1}, lie in G_1 and q_i, q_{i+1} lie in G_2, then the crossing is in the spine between p_{i-1} and p_i; now we exchange the roles of G_1 and G_2 and proceed along the spine.

Note that a crossing only occurs if Q crosses F, but not necessarily, for in the first case above Q crosses F twice and the caterpillar crosses F only once. Each caterpillar must cross F at least once in the above situations. So, for any position sequence S, we have $\alpha(S,Q)\geq\alpha(P,Q)$.

Combined with $\ell(T',Q')\geq\ell(P',Q')$ we get $\ell(T,Q)\geq\ell(P,Q)$, contradicting our assumption that $R(T,Q)$ was a shorter caterpillar that $R(P,Q)$. □

5. Retracts of hypercubes

A *retract* of a graph G is an isometric subgraph H of G such that there is a distance decreasing map of G onto H, which restricted to H is the identity. Bandelt [Ba 84] proved that the median graphs are precisely the retracts of hypercubes (for further references on retracts see [Ba 84] or [CGS 89]). This result also can be proved using expansions. We only sketch that here using the notation introduced above.

We define an *extremal colour* of a median graph G to be a colour F such that, say, $G_1=G_0$. Then G_1 is an *extremal subgraph*. In a tree the end vertices are the extremal subgraphs, and in an n-dimensional hypercube (n–cube, for short) the $(n-1)$–subcubes are the extremal subgraphs. Note that the edges on a geodesic in a median graph all have different colours.

Lemma 7. Let G be a median graph with split G_1,G_2. Then G_1 as well as G_2 contain an extremal subgraph.

Proof. Assume that $G_1\neq G_{10}$. Let x be a vertex in G_1-G_{10} adjacent to a vertex y in G_{10}, and let z be the neighbour of y in G_{20}. Recall that z is the gate for

y in G_2. Let A be the colour of xy, and let F be the colour of yz (i.e. the colour between G_1 and G_2). We will show that colour A does not occur in G_2. Note that, if A occurs in G_{10}, then it occurs in G_{20} as well.

Assume the contrary, and let pq be an edge of A in G_2 with, say, $d(y,p)+1=d(y,q)$. Then u and p are on one side of A, so that x and q are on the other side. Let $P=y\to z\to\ldots\to p$ be a y,p-geodesic. Then there is an x,q-geodesic $Q=x\to t\to\ldots\to q$ with t adjacent to z. Then xt and yz have the same colour, so xt is in F. This implies that x is in G_{10} contradicting the choice of x. So the colour A is fully contained in G_1, and $G_{10}\cup G_2$ is on one side of A and x on the other side.

Repeating this argument, if necessary, we arrive at an extremal subgraph of G fully contained in G_1. Similarly there is an extremal subgraph contained in G_2. □

Using Theorem 3.2.7 from [Mu 80b], we can easily verify that a retract of a hypercube is a median graph. To prove that each median graph can be realized as a retract of a hypercube we use induction on the number of colours.

Let G be a median graph, and let F be an extremal colour with extremal subgraph $G_1=G_{10}$. We embed G in an n-cube Q as in Theorem 3. Then F splits Q into two $(n-1)$-cubes Q_1 and Q_2 with G_i in Q_i, $i=1,2$. By induction there is a retraction of Q_2 onto G_2. Apply the corresponding retraction on Q_1. Then it maps Q_1 onto a copy H_1 of G_2 matched isomorphically via F to G_2. This map preserves G_1. Now we only have to map H_1–G_1 into G_2 in the right way. If u_1 in H_1–G_1 has neighbour u_2 in G_2–G_{20}, then we map u_1 on a neighbour of u_2, which is nearer to G_{20} than u_2. This is possible whenever we have a distance decreasing map of G_2 into itself, which preserves G_{20} and maps vertices of G_2–G_{20} on neighbours nearer to G_{20}.

The existence of such a map can again be proved by induction on the number of colours. We omit the details here.

Actually this is precisely the way how WILKEIT [Wi 86] proved that the so-called quasi-median graphs are the retracts of the Cartesian products of arbitrary complete graphs (see Section 4).

6. Crossing splits

Two splits G_1, G_2 and H_1, H_2 of a median graph G, or their associated colours, are said to be *crossing* if $G_i\cap H_j\neq\emptyset$, for $i,j = 1,2$. Note that, for a split G_1, G_2 of G, the subgraph G_1 is extremal if and only if each colour occurring in G_1 crosses F_{12} (see [Mu 90]). We use this fact in the following theorem, which has a very simple proof due to the expansion technique.

Theorem 8. Let G be a median graph. Then G contains n pairwise crossing splits if and only if G contains an n-cube as an induced subgraph.

Proof. If G contains an n-cube, then the n colours of this cube extend to pairwise crossing splits in G because of the 4-cycle property.

Assume G contains n pairwise crossing splits G_1^k, G_2^k for $k=1,\dots,n$. Without loss of generality, we may assume that G has no other spilts. Otherwise we could contract these, and the contraction would still contain n pairwise crossing splits, and the existence of an n-cube in contraction yields an n-cube in any expansion by its history.

Note that now every colourhalf G_i^k is an extremal subgraph of G, i.e., for $k=1,\dots,n$, colour F_{12}^k yields an isomorphism between $G_1^k=G_{10}^k$ and $G_2^k=G_{20}^k$. Using induction on the number of colours n in G, we may conclude that both G_1^k and G_2^k are $(n-1)$-cubes, so that G is an n-cube. □

7. The hull number of a median graph

The intersection of convex sets in a graph is again convex. This gives rise to the following definition. Let W be a subset of vertices in a graph $G=(V,E)$. The *convex hull* of W, denoted by $Con(W)$, is the smallest convex subgraph of G containing W (see [Mu 80b], where it was termed the convex closure). A set $S \subseteq V$ *generates* G if $Con(S)=G$. In [ES 85] EVERETT and SEIDMAN introduced the *hull number* $h(G)$ of a graph G to be the size of a *minimum generating set*. Here of course, minimum means that there is no generating set with fewer vertices.

Any two diametrical vertices (vertices at largest distance) generate a hypercube. So $h(Q)=2$, for any hypercube Q except K_1. In a tree T we need all end vertices to generate T. By convention an end vertex will be the vertex of degree zero if $T=K_1$, and a vertex of degree one otherwise. Clearly, $h(T)$ is the number of end vertices in T.

In this subsection we consider (minimum) generating sets of median graphs. We say that a set W *touches* a subgraph H of G if H contains a vertex of W. The following three results are obvious (we use the above notations).

Lemma 9. If S generates the median graph G and G' is a contraction of G, then S' generates G'.

Corollary 10. If G' is a contraction of the median graph G, then $h(G) \geq h(G')$.

Lemma 11. If G_1,G_2 is a split in a median graph G generated by S, then S touches G_1 as well as G_2.

The main result of this subsection is the following theorem.

Theorem 12. Let S be a set of vertices touching each extremal subgraph of a median graph G. Then S generates G.

Proof. We use induction on the number of expansions. Let F be an extremal colour with split G_1,G_2 and $G_1 = G_{10}$. We may take G_2 as the contraction of G with respect to F. Note that every colour in G_1 occurs in G_{20} as well, and vice versa.

Every extremal colour of G distinct from F is an extremal colour of G_2. So all extremal subgraphs of G_2 associated with these colours are touched by S'. If A is a non-extremal colour in G_1, then it is also non-extremal in G_{20} as well as in G_2.

Assume that B is an extremal colour in G_2 that is not extremal in G. Then G_{20} must be contained in the extremal subgraph of B. Since S touches G_1, it follows that S' touches G_{20}, so it touches the extremal subgraph of B in G_2 as well. Hence S' touches all extremal subgraphs of G_2.

By induction S' generates G_2. Let $w_1,x_1,\ldots,z_1$ be the vertices of S in G_1, and let $w_2,x_2,\ldots,z_2$ be their respective neighbouring gates in G_{20}. Since S generates G, it touches G_2, say in v. Then w_2 lies in $I(w_1,v)$, etcetera. So $Con(S)$ contains $w_2,x_2,\ldots,z_2$. Therefore $Con(S)$ contains $Con(S') = G_2$, in particular $Con(S)$ contains G_{20}. Take any vertex p_2 in G_{20} with neighbouring gate p_1 in G_1. Then $I(w_1,p_2)$ contains p_1. So $G_1 = G_{10}$ is contained in $Con(S)$ as well, and we are done. □

The following theorem is an immediate consequence.

Theorem 13. Let G be a median graph. Then $h(G)$ is equal to the minimum number of vertices touching all extremal subgraphs of G.

It is easily seen that one can actually decrease the hull number by contractions. But what are the contractions that preserve the hull number? In a tree one can contract all internal edges, thus obtaining a *star* (a $K_{1,n}$) with the same number of end vertices. Contracting any further edge decreases the hull number. In a hypercube we can contract all colours but one, thus obtaining the star $K_{1,1}$ with the same hull number. By convention we will consider K_1 also to be a star.

A *star contraction* of a median graph G is a star obtained by successive contractions of G. Let T be a star contraction of G with the maximum possible number of end vertices. We define $\tau(G)$ to be the number of end vertices of this star T. Then we get the following problem.

Question. For which median graphs G do we have $h(G) = \tau(G)$?

8. Quasimedian graphs

Almost all of the above results can be generalized to ***quasimedian graphs***, which generalize median graphs. These graphs were introduced and characterized by another expansion procedure in [Mu80b]. For the relevant theorems on retracts see [CGS 89] and [Wi 86], and for the generalization of the dynamic search problem, see [CGS 89].

References

[Av 61] S.P. Avann, Metric ternary distributive semi-lattices, Proc. Amer. Math. Soc. **12** (1961) 407–414.

[Ba 84] H.-J. Bandelt, Retracts of hypercubes, J. Graph Theory **8** (1984) 501–510.

[BB 84] H.-J. Bandelt and J.P. Barthélemy, Medians in median graphs, Discrete Appl. Math. **8** (1984) 131–142.

[BH 83] H.-J. Bandelt and J. Hedlíková, Median algebras, Discrete Math. **45** (1983) 1–30.

[CGS 87] F.R.K. Chung, R.L. Graham and M.E. Saks, Dynamic search in graphs, in Discrete Algorithms and Complexity, Academic Press (1987) 351–388.

[CGS 89] F.R.K. Chung, R.L. Graham and M.E. Saks, A dynamic location problem for graphs, Combinatorica **9** (1989) 111–131.

[Dr 87] H. DRAŠKOVIČOVÁ, Weak direct product decompositions of algebras, Contributions to General Algebra **4**, Hölder–Pichler–Tempsky (1987), Wien.

[Es 85] M.G. EVERETT and S.B. SEIDMAN, The hull number of a graph, Discrete Math. **57** (1985) 217–223.

[Is 80] J.R. ISBELL, Median algebra, Trans. Amer. Math. Soc. **260** (1980) 319–362.

[JS] P.K. JHA and G. SLUTZKI, Convex-expansions algorithms for recognition and isometric embedding of median graphs, Ars Comb. **34** (1992) 75–92.

[vM 77] Supercompactness and Wallman spaces, Ph. D. Thesis, Vrije Universiteit Amsterdam, 1977.

[MMR 94] F.R. McMorris, H.M. Mulder and F.R. Roberts, The median procedure on median graphs, Report 9413/B, Econometrisch Instituut EUR, 1994, submitted.

[Mu 78] H.M. MULDER, The structure of median graphs, Discrete Math. **24** (1978) 197–204.

[Mu 80a] H.M. MULDER, n-cubes and median graphs, J. Graph Theory **4** (1980) 107–110.

[Mu 80b] H.M. MULDER, The interval function of a graph, Math. Centre Tract 132, Math. Centre Amsterdam (1980).

[Mu 90] H.M. MULDER, The expansion procedure for graphs, in: R. Bodendiek ed., Contemprorary methods in graph theory, Wissenschaftsverlag, Mannheim etc., 1990, pp. 459–477.

[Ne 71] L. NEBESKÝ, Median graphs, Comment. Math. Univ. Carolinae **12** (1971) 317–325.

[Wi 86] E. WILKEIT, Isometrische Untergraphen von Hamming–Graphen, Doctoral thesis, Universität Oldenburg (1986).

[Wr 87] P. WINKLER, The metric structure of graphs: theory and applications, in: Surveys in Combinatorics 1987, C. Whitehead ed., Cambridge University Press (1987) 197–221.