COMBINATORICA 5 (1) (1985) 67—79

AN ERDŐS—GALLAI CONJECTURE

L. PYBER

Received 27 March 1984

The following conjecture of P. Erdős and T. Gallai is confirmed: every graph on n vertices can be covered by $n-1$ circuits and edges.

0. Introduction

P. Erdős and T. Gallai posed a conjecture in 1966 [3], this conjecture was also mentioned as an unsolved problem in the book of Bondy and Murty [1]. In 1982 Tao Mao-qi, Shen Yun-qiu and Ku Tung-shin proved the conjecture in a slightly stronger form for planar graphs [8]. What they prove is the following:

Theorem 0. *Let G be a 2-edge-connected planar graph of n vertices. Then G can be covered by $(n-2)$ circuits.*

Their proof is somewhat lengthy, here we give a shorter one. Our main result is

Theorem 1. *A graph G of n vertices can be covered by $n-1$ circuits and edges.*

Notice that Theorem 1 is sharp if G is a tree.

We also prove the following stronger version of Theorem 1:

Theorem 2. *Let G be a graph of n vertices and $\{C_1, \dots, C_k\}$ be a set of circuits and edges such that $\bigcup_{i=1}^{k} E(C_i)=E(G)$ and k is minimal. Then we can choose k different edges, $e_i \in E(C_i)$, such that these edges form a forest in G.*

Finally we prove another theorem for graphs not containing C_4 (circuit of length four), which is stronger in some sense.

Theorem 3. *Let G be a graph of n vertices not containing C_4. Then G can be covered by $[(n-1)/2]$ circuits and $n-1$ edges.*

Basic to our proofs is the following theorem of Lovász [4]:

AMS subject classification (1980): 05 C 38

Theorem 4. *A graph of n vertices can be covered by* $\leq [n/2]$ *disjoint paths and circuits.*

Notation. $d_G(x)$ denotes the degree of the vertex $x \in V(G)$ in the graph G. $d(x, y)$ denotes the distance of the vertices x and y in G. For a subgraph H of G, G/H and $G \setminus H$ denote the graphs obtained from G by contracting or deleting the edges of H. $\langle D \rangle$ denotes the subgraph of G induced by the set $D \subset V(G)$.

1. The proof of Theorem 0

We use the following wellknown lemma (see e.g. [6]):

Lemma 1.1. *Let G be a planar graph, then it has an edge contained in at most two triangles of G.* ∎

Lemma 1.2. *If G is an arbitrary graph* $e=(x, y) \in E(G)$ *and* G/e *has a cut-edge then either*

(a) *G has a cutvertex or*

(b) *there is a vertex* $z \in V(G)$ *such that the only neighbours of z in G are x and y.*

Proof. If $(p, q) \in E(G/e)$ is a cut-edge in G/e then p and q are cutvertices of G/e or components of $(G/e) \setminus (p, q)$. If p and q are both cutvertices then they are the contracted images of two cutsets in G, say of P and Q. We contracted only one edge to obtain G/e so one of the two sets has only one vertex, i.e. G has a cutvertex.

If p is a component of $(G/e) \setminus (p, q)$ and G has no cutvertices then p is the image of a vertex $z \in V(G)$ with $d_G(z) \leq 2$. Here, in fact, $d_G(z)=2$ and (b) follows. ∎

Now we prove Theorem 0 by induction on $n=|V(G)|$ for planar graphs.

Lemma 1.3. *If the graph G has a cutvertex and Theorem 0 holds for all graphs H with* $|V(H)|<|V(G)|$ *then it holds for G.*

Proof. Let p be a cutvertex of G. Deleting p we obtain components $K_1, \ldots, K_r$ $(r \geq 2)$. The graphs $\langle V(K_i) \cup p \rangle$ are 2 edge-connected, i.e. these graphs can be covered by $n_i - 2$ circuits, respectively, here $n_i = |V(K_i)|+1$. The set of all these circuits covers G, $\sum_{i=1}^{r} (n_i - 1) + 1 = |V(G)| = n$ and this leads to $\sum_{i=1}^{r} (n_i - 2) \leq n-2$. ∎

Let $(x, y)=e$ be an edge of G contained in at most two triangles (exists by Lemma 1.1). If G has no cutvertex and G/e has a cut-edge then we can apply Lemma 1.2. Indeed, (z, x) and (z, y) are contained in at most one triangle. If $G/(z, x)$ and $G/(z, y)$ are not 2-edge-connected then by Lemma 1.2 G is a triangle. Here $n=3$, in this case the Theorem is trivially true.

It suffices to prove the induction hypothesis for 2-connected graphs G, having an edge $e=(x, y)$ contained in at most two triangles, such that G/e is 2-edge-connected.

We distinguish three cases according to the number of triangles containing e. We always have $n-3$ circuits $\{C_1', \ldots, C_{n-3}'\}$ covering G/e.

Case 0. If there are no triangles on e, then the contraction is injective on $E(G)\setminus e$. The circuits covering G/e correspond to circuits of G in the natural way, such that the latter $n-3$ circuits cover $E(G)\setminus e$. G is 2-edge-connected, therefore we have an $(n-2)$-th circuit C in G covering e.

Case 1. Suppose that there is exactly one triangle containing $e=(x, y)$, say $T=(x, y, z)$. Again, we have $n-3$ circuits C_i which correspond to the circuits C_i' of G/e (the contracted images of the circuits C_i are the C_i'). Indeed, the circuits C_i cover $G\setminus T$. Adding circuit T we obtain an $(n-2)$-element covering of the graph G.

Case 2. Suppose that (x, y) is contained in exactly two triangles, namely, in (x, y, z) and (x, y, v). The contracted image of x and y in G/e is denoted by p, the image of z by z'. As $(p, z')\in E(G/e)$, we might suppose that $(p, z')\in E(C_1')$. Let us extend the circuits C_i' to circuits C_i of G arbitrarily. We define the circuit C_0 in the following way: If $(x, y)\in E(C_1)$ then $C_0=C_1$. If not, then one of the vertices x and y, say x, is not in C_1, for the image of C_1 in G/e is a circuit. We have $(y, z)\in E(C_1)$, for $(p, z')\in E(C_1')$. Now let us define C_0 by $E(C_0)=(E(C_1)\setminus(y, z))\cup\{(x, y), (x, z)\}$. Let C be the circuit (x, z, y, v). It is easy to see that the circuits $C, C_0, C_1, \ldots, C_{n-3}$ cover G.

The proof of Theorem 0 is now complete. ∎

2. Preliminaries to the proof of Theorem 1

Definition. Let G be a connected graph, $a\in V(G)$. We define the *distance classes* $D_i^a=\{x|x\in V(G), d_G(x, a)=i\}$ for $i=0, \ldots, r$. (For the sake of simplicity we omit the upper index.)

Lemma 2.1. *(a) $D_0, \ldots, D_r$ is a partition of $V(G)$.*

(b) For any $x\in D_i$ there exists an $y\in D_{i-1}$ with $(x, y)\in E(G)$. Choosing such an edge (x_j, y_j) for all $x_j\in D_i$ we obtain a star system W_i in G, i.e. a subgraph which is the union of disjoint stars. The centre of a star in W_i is in D_{i-1}.

(c) The union of the star systems W_i is a tree W of G which we call a distance tree of G.

(d) Let $G_{r+1}=\langle D_r\rangle$ and $G_i=\langle D_i\cup D_{i-1}\rangle\setminus\langle D_i\rangle$ for $i=1, \ldots, r$. Then $W_i\subset G_i$ and $E(G_1), \ldots, E(G_{r+1})$ is a partition of $E(G)$.

(e) If P is a path in G such that its vertices are in $\bigcup_{j\geqq i} D_j$ and its endvertices are in D_i, then there exists a circuit C of G containing P. Such paths are called embeddable.

Proof. Statements (a)—(d) are wellknown or trivial.

(e) Let $x_1, x_2\in D_i$ be the endvertices of P. We have two unique paths $P_1, P_2\subset W$ leading from D_0 to x_1 and x_2 respectively. Indeed, P_1 and P_2 have one common vertex with D_j for $j\leqq i$ and none for $j>i$. P_1 and P_2 have a common vertex, D_0. They have a last common vertex $v\in D_j, j<i$, where j is the largest index for which such a v exists. The union of P and the subpaths of P_1, P_2 leading from v to x_1, x_2 resp. form a circuit C in G. ∎

The above lemma contains the basic ideas of the proof of Theorem 1. Applying Theorem 4 for the graphs G_i we get strong bounds for the number of covering edges and circuits of G.

However, to prove Theorem 1 we also have to consider the proof of Theorem 4, and we must choose W in a special way: Let us denote the vertices in D_i by $\{a_t^i\}$ $(t=1, ..., |D_i|)$. For any $a_t^i \in D_i$ we have an $a_s^{i-1} \in D_{i-1}$ such that $e_t^i=(a_t^i, a_s^{i-1}) \in E(G)$ and s is minimal. We define W_i as the set of the edges e_t^i and W as the union of the star systems W_i. (We fix a and the order of the elements of the D_i-s arbitrarily and in the proof of Theorem 1 W will denote the distance tree of G defined by this order in the above "greedy" way.)

The original proof of Theorem 4 contained a minor error. It was corrected by A. Donald therefore we refer to his paper [2].

Notations. We denote the set of neighbours of $x \in V(G)$ in the graph G by $N(x)$. If Σ is a covering of G by disjoint paths and circuits, $x \in V(G)$, then $t(x, \Sigma)$ denotes the number of paths in Σ with endvertex x. We allow paths of length 0, they increase $t(x, \Sigma)$ in the corresponding vertex x by 2.

Lemma 2.2. *Suppose x is a vertex of a graph G, $A=\{a_1, ..., a_k\} \subset N(x)$, $(k \geqq 1)$, G' is the graph obtained from G by deleting the edges between x and A. Let Σ' be a covering of G' by disjoint paths and circuits, such that $t(p, \Sigma') \geqq 1$ for $p \in N(x)$. Then there exists a covering Σ of G by disjoint paths and circuits such that $|\Sigma|=|\Sigma'|$, and*

$$\text{for} \quad p \in V(G) \setminus (A \cup x) \quad t(p, \Sigma) = t(p, \Sigma'),$$

$$\text{for} \quad p \in A \quad t(p, \Sigma) = t(p, \Sigma')-1 \quad \text{and} \quad t(x, \Sigma) \geqq t(x, \Sigma').$$

This lemma is essentially a more detailed version of Lemma 7 in [2]. Here we repeat the main steps of the proof in (6).

Proof. For each $i=1, ..., k$, we define a sequence of vertices of $N(x)$: $a_{i,0}, ..., a_{i,r_i}$ by recursion as follows. Let $a_{i,0}=a_i$. Suppose $a_{i,\mu}$ is defined; since $a_{i,\mu} \in N(x)$ there is a path $U_{i,\mu}$ of Σ' that has $a_{i,\mu}$ as an endvertex. If $U_{i,\mu}$ does not pass through x, end the sequence; if x lies on $U_{i,\mu}$, trace the path from $a_{i,\mu}$ to x and let $a_{i,\mu+1}$ be the last vertex on the path before it reaches x.

Observations. $a_{i,\mu}$ is adjacent to x in G' iff $\mu \geqq 1$.

If $a_{i,\mu}=a_{j,\nu}$, then $i=j$ and $\mu=\nu$. For suppose $a_{i,\mu}=a_{j,\nu}$ and $\mu \leqq \nu$. If $\mu=0$, then $a_{i,\mu}=a_i$, so $\nu=0$ and $i=j$. If $\mu \geqq 1$, then $a_{i,\mu}$ lies on $U_{i,\mu-1}$, which has an endvertex at $a_{i,\mu-1}$ and contains the edge $(x, a_{i,\mu})$. Similarly, $a_{i,\mu}=a_{j,\nu}$ lies on $U_{j,\nu-1}$ which contains $(x, a_{j,\nu})=(x, a_{i,\mu})$. Since the paths of Σ' are edge disjoint, $U_{i,\mu-1}=U_{j,\nu-1}$ and $a_{i,\mu-1}=a_{j,\nu-1}$. Repeating this argument we get $a_{i,0}=$ $=a_{j,\nu-\mu}$ and $\nu-\mu=0$ so $\mu=\nu$ and $i=j$.

This implies that the above sequences are finite.

The edge $(x, a_{i,\mu})$ lies on $U_{i,\mu-1}$ if $\mu \geqq 1$.

We now define a function f that maps elements of Σ' into disjoint paths and circuits covering G.

Let $U \in \Sigma'$, if it is not a $U_{i,\mu}$, then $f(U)=U$. Otherwise we distinguish two sets of mutually exclusive cases:

(a) $U=U_{i,\mu}$ contains x.

(a′) $U=U_{i,\mu}$ does not contain x.

(b) $U=U_{i,\mu}=U_{j,\nu}$ where $(i, \mu) \neq (j, \nu)$.

(b′) $U=U_{i,\mu}$ and the labelling is unique.

We define $f(U)$ for the four possible combinations.

If (a) and (b) hold we delete the edges $(x, a_{i,\mu+1})$ and $(x, a_{j,\nu+1})$, add the edges $(x, a_{i,\mu})$ and $(x, a_{j,\nu})$ to obtain $f(U)$ from U.

If (a) and (b′) hold we delete $(x, a_{i,\mu+1})$ and add $(x, a_{i,\mu})$.

If (a′) and (b) hold we add the edges $(x, a_{i,\mu})$ and $(x, a_{j,\nu})$.

If (a′) and (b′) hold we add $(x, a_{i,\mu})$.

Remark. If U is a path of length 0 then (a′) and (b′) hold, therefore $f(U)$ is an edge of G.

We now show that the paths and circuits formed by f are disjoint, and cover G. If (y, z) is not of the form $(x, a_{i,\mu})$ then it appears in some U of Σ' and also in, and only in, $f(U)$. Examination of the cases shows that $(x, a_{i,\mu})$ appears in $f(U_{i,\mu})$, so it remains to prove only that if $(x, a_{k,\varrho})$ is in $f(U_{i,\mu})$, then $U_{k,\varrho}=U_{i,\mu}$.

If (b) holds for $U=U_{i,\mu}$, then $U_{i,\mu}=U_{j,\nu}$ where $(i, \mu)\neq(j, \nu)$. The only edges of $f(U_{i,\mu})$ incident to x are $(x, a_{i,\mu})$ and $(x, a_{j,\nu})$. Since $a_{k,\varrho}\neq a_{i,\mu}$, by assumption, it follows that $a_{k,\varrho}=a_{j,\nu}$ so $(k, \varrho)=(j, \nu)$. If (a) and (b′) hold for $U=U_{i,\mu}$, let $U_{i,\mu}=(a_{i,\mu}, \dots, a_{i,\mu+1}, x, y, \dots, z)$. If $(x, a_{k,\varrho})$ is in $f(U_{i,\mu})$ and $a_{k,\varrho}\neq a_{i,\mu}$ then $a_{k,\varrho}=y$. Thus $z=a_{k,\varrho-1}$ and $U_{i,\mu}=U_{k,\varrho-1}$. But $U_{i,\mu}$ is uniquely labeled so $(i, \mu)= =(k, \varrho-1)$ and $a_{i,\mu}=a_{k,\varrho-1}=z$. Thus $U_{i,\mu}$ has to be a circuit, a contradiction. If (a′) and (b′) hold for $U_{i,\mu}$, the only edge of $f(U_{i,\mu})$ incident to x is $(x, a_{i,\mu})$.

Thus $\Sigma=f(\Sigma')$ is a covering of G by disjoint paths and circuits and $|\Sigma|= =|\Sigma'|$.

Let us verify now our additional observations. If $p\in V(G)$ is not of the form $a_{i,\mu}$ then the number of paths with endvertex p is indeed the same for Σ and Σ'. If $p=a_i$ then p is an endvertex of $U_{i,0}$ but not of $f(U_{i,0})$ and for $U\neq U_{i,0}$, p is an endvertex of U iff it is an endvertex of $f(U)$. If $p=a_{i,\mu}$, $\mu\neq 0$ then p is an endvertex of $U_{i,\mu}$ and of $f(U_{i,\mu-1})$ but not of $f(U_{i,\mu})$ and $U_{i,\mu-1}$, for any other U, p is an endvertex of U iff it is an endvertex of $f(U)$. (If U is a path of length 0 then the situation is different, here $p=U$ might become a single endvertex of $f(U)$ from a "double" endvertex of U according to the above cases.) If x is an endvertex of U then it is an endvertex of $f(U)$.

Throughout the proof we never used the assumption that the paths have edges so it also works if there are paths of length 0 in Σ.

The proof is complete. ∎

Remark 2.3. Suppose that the conditions of Lemma 2.2 are not satisfied, there is a vertex in $N(x)\setminus A$, say b, with $t(b, \Sigma')=0$. (For the sake of simplicity we suppose the existence of exactly one such vertex.) Adding the "path" b to Σ' and performing the construction of the above proof we obtain a covering Σ_1 of G with $|\Sigma_1|=|\Sigma'|+1$.

We have two possibilities: either $f(b)=b$ then we simply did not need b in constructing $\Sigma=\Sigma_1\setminus b$, a covering of G with the same properties as in Lemma 2.2, or we needed b and $f(b)$ is the edge (b, x).

In the latter case $\Sigma=\Sigma_1\setminus f(b)$ is a covering of $G\setminus(b, x)$ with $|\Sigma|=|\Sigma'|$ and the statements of Lemma 2.2 are valid for Σ with the exception that $t(b, \Sigma)= =t(b, \Sigma')+1$. (It easily follows from the proof of Lemma 2.2.) ∎

Definition. A vertex of a graph is called *odd* or *even* if it is of odd or even degree, respectively. A graph G is called *odd* if it has at most one even vertex.

Theorem 4 is true in a stronger form for odd graphs:

Corollary 2.4. *An odd graph of n vertices can be covered by* $[n/2]$ *disjoint paths such that each odd vertex of G is the endvertex of exactly one covering path.*

Let us recall the following ([6]):

Lemma 2.5. *Let G be a connected graph, F a spanning tree of G and p an arbitrary vertex of G. There exists a forest* $E\subset F$, *such that* $G'=G\setminus E$ *is an odd graph and if* $|V(G)|$ *is odd then p is the even vertex of* G'.

Proof. Let F be a spanning tree of G and $x_1, \ldots, x_k$ the even vertices of G. Let P_i denote the unique path between p and x_i in F. Set $E=\sum_{i=1}^{k} E(P_i) \pmod 2$. For each $x\in V(G)$ we have $d_E(x)=\sum_{i=1}^{k} d_{P_i}(x) \pmod 2$. Cosequently for $x\in V(G)$, $x\neq p$, $d_E(x)$ is odd iff $d_G(x)$ is even. With $G'=G\setminus E$ the lemma follows. ∎

Remark 2.6. If Theorem 1 holds for all 2-connected graphs then it follows for any graph G by an easy induction on $n=V(G)$. (Like in the proof of Theorem 0.)

From now on we suppose that G is 2-connected. It is wellknown that in such a graph any two edges are contained by a circuit.

3. The proof of Theorem 1

It seems to be impossible to present the proof as a series of independent lemmas, therefore we start with

3.0. Sketch of the proof

We have a 2-connected graph G with a fixed "greedy" distance tree W (for definitions see Lemma 2.1). We want to cover most of the edges by paths embeddable into circuits and put the rest pairwise into circuits.

The first step is the covering of G_{r+1}. By Lemma 2.5 it has a large subgraph which can be covered by a few paths such that almost all points in D_r are endvertices of only one of the paths (see 2.4). This enables us to extend the the constructed paths to cover a star system of G_r and remain embeddable. Still we have an uncovered forest from G_{r+1}. This we add to $G_r\setminus W_r$ and try to cover the resulting graph.

We have to perform the so called general step. Again we cover a large part of the edges of the above graph such that most of the points in $D_r\cup D_{r-1}$ are endvertices of exactly one path. Now we apply Lemma 2.2 to obtain a new set of paths and circuits which cover also W_r such that the endvertices of the paths are in D_{r-1} (i.e. they are embeddeble). Again most of the points in D_{r-1} are the endvertices of some path. The extensions of these paths cover a star system in G_{r-1} (which we define later). There remains an uncovered forest on $D_r\cup D_{r-1}$. A star system of it can be covered by the extensions of paths in G_{r+1}. We contract this star system in the forest and add the contracted image to $G_{r-1}\setminus W_{r-1}$. A covering of the resulting graph can be extended into a set of circuits in G which cover the uncontracted edges of the forest.

We proceed by performing the general step inductively.

A final calculation shows that we did not use too many circuits.

3.1. *The covering of* G_{r+1}

We denote the connected components of G_{r+1} by $K_1, \ldots, K_f$. Let us choose one vertex p_j in each K_j with $|V(K_j)|$ odd. By Lemma 2.5 we have the forests $E_j\subset K_j$ $(j=1, \ldots, f)$ such that the graphs $K_j'=K_j\setminus E_j$ are odd and their only even vertices are p_j. Applying Corollary 2.4 we might cover K_j' by $[|V(K_j)|/2]$ paths such that the only vertex of K_j' which is not an endvertex of a covering path is p_j (if it exists). Set $E=\bigcup_{j=1}^{f} E_j$. It is indeed a forest on the points of D_r. Let α_{r+1} denote the number of p_j-s (i.e. the number of odd $|V(K_j)|$-s). The number of the considered covering paths is

$$\mathcal{K}_{r+1} = \sum_{j=1}^{f}\left[\frac{|V(K_j')|}{2}\right] = \frac{|D_r|}{2}-\frac{\alpha_{r+1}}{2}.$$

Let us define the graph $H_r=(G\setminus G_{r+1})\cup E$.

We can summarise the above results as follows:

3.1a) $\langle D_r\rangle$ is a forest in $H_r\subset G$.

3.1b) $G\setminus H_r$ can be covered by $\mathcal{K}_{r+1}$ paths such that we have only α_{r+1} points in D_r which are not endvertices of some path.

3.1c) All covering paths are embeddable (see Lemma 2.1e). ▮

In the general step we start with a graph H_i, having similar properties as H_r, and obtain H_{i-1}.

3.2. *The general step*

The graph H_i is defined on $\bigcup_{j=0}^{i} D_j$. The subgraph of H_i induced by D_i is a forest E. (If possible we omit the index i, for example we write E instead of E_i in the $(r+1-i)$-th general step.) $H_i\setminus E=\langle \bigcup_{j=0}^{i} D_j\rangle\setminus\langle D_i\rangle$ where the latter graphs are induced subgraphs of G.

We also have some paths on $D_{i+1}\cup D_i$ (constructed in the $(r-i)$-th step) such that there are at most α_{i+1} points in D_i which are not endvertices of some path. (α_{r+1} is defined in 3.1, α_i will be defined later for $i\leqq r$.)

3.2.1. *The isolated vertices of* $G_i\setminus W_i$

Let us consider the graph $G_i\setminus W_i$. I_i denotes the set of isolated vertices of this graph in D_i. For later purpose we want to "get rid of" I_i.

Suppose we have an edge (x, y) of E with $x\in I_i$. We contract this edge and identify $V(E/(x, y))$ with $V(E)\setminus x$ in the natural way. Applying a series of such contractions we finally obtain a forest E^* on $D_i\setminus I_i$. Indeed we contracted at most $|I_i|$ edges. We denote their set by S^*. The contraction of an edge in a forest is injective on all other edges and the graph obtained by the contraction is also a forest. Consequently S^* corresponds to a set of edges S in E with $|S|=|S^*|$. Moreover $E^*=E/S$ and the contraction of S is injective on the set of edges of E which are not in S.

Let $H_i^*=(H_i\setminus E)\cup E^*$. We define $R\subset H_i^*$ by $R=(G_i\setminus W_i)\cup E^*$. Then we have:

(a) $|V(H_i)|-|V(H_i^*)|=|S|=\gamma_i\leqq|I_i|$.

(b) If Σ is a covering set of circuits for H_i^*, then we can extend the circuits in such a way that the extended circuits cover $E(H_i)\setminus S$. (Actually we need a much more difficult statement.)

(c) For each vertex x of H_i^* which is a point of D_i, we have an edge $(x, y)\in E(H_i\setminus W_i)$ with $y\in D_{i-1}$. (That is going to be important.)

3.2.2. The covering of the main part of R

As in 3.1 let us consider the connected components K_j of R $(j=1, \dots, f)$. If K_j has a vertex in D_i then it also has an edge connecting this vertex to some vertex in D_{i-1}. Choosing one such edge for each vertex in $D_i\cap V(K_j)$, we obtain a star system which can be embedded in some spanning tree F_j of K_j. If $V(K_j)\subset D_{i-1}$ then we choose the spanning tree F_j arbitrarily.

The number of K_j-s having at least one vertex in D_i and $|V(K_j)|$ odd is denoted by β_i. For each such K_j let us choose a vertex $p_j\in V(K_j)\cap D_i$, in other components K_j with $|V(K_j)|$ odd we choose p_j arbitrarily. The number of components with $V(K_j)\cap D_{i-1}$ and $|V(K_j)|$ odd is denoted by α_i. (We completed the definition of the important numbers $\alpha_i, \beta_i, \gamma_i$.) By Lemma 2.5 we have the forest $E_j\subset F_j$, $(j=1, \dots, f)$, such that $K_j'=K_j\setminus E_j$ is an odd graph and if it has an even vertex then it is p_j. Applying Corollary 2.4 we might cover the graph K_j' by $[|V(K_j')|/2]$ disjoint paths such that any $x\in V(K_j')$, $x\neq p_j$, is an endvertex of exactly one path.

The next step is to change these paths using the method of Lovász.

3.2.3. Covering of W_i'

For $(x, y)\in E(W_i)$ let $x\in D_i$ and $y\in D_{i-1}$. The set of edges (x, y) with $x\notin I_i$ is denoted by W_i'. We associate each $(x, y)\in W_i'$ with the unique (and existing) component K_j containing x. The edges of W_i' associated with a certain K_j form q stars $S_1, \dots, S_q$ with their centers say $a_{t_1}^{i-1}, \dots, a_{t_q}^{i-1}$ in D_{i-1}, where $t_1>t_2>\dots>t_q$. (For notation see the definition of the "greedy" distance tree W in part 2.) We fixed j, Σ_0 denotes the covering set of paths for K_j'. Let us first apply Lemma 2.1 to the graph $K_j'\cup S_1$ with $\Sigma'=\Sigma_0$. We have at most one vertex which is not an endvertex of a path: p_j. The following cases might occur:

(a) p_j is not a neighbour of $a_{t_1}^{i-1}$ in $K_j'\cup S_1$. By Lemma 2.2 we have a covering set of disjoint paths and circuits Σ_1 for $K_j'\cup S_1$ such that the vertices v in $V(S_1)\setminus a_{t_1}^{i-1}$ are the only ones with $t(v, \Sigma_1)<t(v, \Sigma_0)$. These are no endvertices of any path in Σ_1.

(b) $e=(a_{t_1}^{i-1}, p_j)\in E(S_1)$, then we obtain a covering Σ_1 of $K_j'\cup(S_1\setminus e)$ with the same properties as in (a), except that $t(p_j, \Sigma_1)=t(p_j, \Sigma_0)=0$

(c) $e=(a_{t_1}^{i-1}, p_j)\in E(K_j')$. In that case we apply Remark 2.3 and obtain a covering Σ_1. If Σ_1 does not cover e (that is, the second possibility of 2.3 occurs) then $t(p_j, \Sigma_1)=t(p_j, \Sigma_0)+1=1$. Now the only vertices of $K_j'\cup S_1$ which are not endvertices of some path of Σ_1 are the vertices in $V(S_1)\setminus a_{t_1}^{i-1}$. If the first possibility occurs then Σ_1 has the same properties as in (a).

In the next step we try to extend Σ_1 to a covering Σ_2 which covers S_2 as well. Here we have cases similar to (a), (b) and (c). We have to note that a neighbour of $a_{t_1}^{i-1}$ in S_1 can not be a neighbour of $a_{t_2}^{i-1}$ in $K_j'\cup S_1\cup S_2$ because of the "greedy" choice of W_i and $t_2<t_1$. We proceed inductively and finally obtain a set of disjoint paths and circuits Σ_q such that:

(1) Σ_q covers the edges of K_j and the associated edges of W_i with at most one exception (if the even vertex $p_j\in K_j'$ exists).

(2) $t(x, \Sigma_q) = 0$ for $x\in V(K_j')\cap D_i$.

(3) $t(x, \Sigma_q) \geqq 1$ for $x\in V(K_j')\cap D_{i-1}$.

Proof. To prove (1) notice that during the q steps case (b) might occur at most once. Moreover in case (c) (in the i-th step) if we have an uncovered edge (uncovered by Σ_i) then p_j becomes an endvertex of a path so (b) or (c) can not occur again. In case (b) (in the r-th step) (c) can not occur later because of the "greedy" choice of W_i. These observations yield (1). Each vertex x in $V(K_j')\cap D_i$ is the endvertex of some edge in W_i' by definition. So among the q steps we have one for each such $x\neq p_j$ after which x is not an endvertex of any path. The above observations concerning the occurrences of cases (b) and (c) show us that (2) is also true for $x=p_j$.

(3) is trivial, for if $x\in V(K_j')\cap D_{i-1}$ then we have $1\leqq t(x, \Sigma_0)\leqq t(x, \Sigma_1)\leqq \dots$ $\dots\leqq t(x, \Sigma_q)$ by Lemma 2.2 and Remark 2.3. Let us summarise our results concerning the K_j'-s and the associated edges of W_i. The new covering sets of paths and circuits and the coverings of the K_j'-s with $V(K_j')\subset D_{i-1}$ together form a set of disjoint paths and circuits such that:

(i) none of the endvertices of the paths are in D_i,

(ii) we have at most β_i edges (see definition at 3.2.2) uncovered in $\left(\bigcup_{j=2}^{f} K_j'\right)\cup$ $\cup W_i$, one for each component with a "troublemaker" even vertex $p_j\in D_i$,

(iii) all the vertices in D_{i-1} are endvertices of some covering path with the exception of α_i vertices of the form p_j. (The p_j-s are in D_{i-1} if $V(K_j')\subset D_{i-1}$.)

3.2.4. Covering of the remaining forest

Let us define a forest F in the following way: We have the forests $E_j=K_j\setminus K_j'$, the star system $W_i\setminus W_i'$ and the trees $K_j\supset F_j\supset E_j$ $(j=1, \dots, f)$. (See definitions at 3.2.2 and 3.2.3 respectively.) By the choice of $W_i\setminus W_i'$ (its vertices form the set of "isolated" vertices I_i) and the choice of the F_j-s the forest $\left(\bigcup_{j=1}^{f} E_j\right)\cup(W_i\setminus W_i')$ might be embedded into a forest $F\subset H_i^*$ which contains at least one edge (x, y)

for any $x\in D_i$, with $y\in D_{i-1}$. Choosing such edges for each $x\in D_i$, we obtain a star system $S\subset F$.

In the former general step we obtained paths and circuits in H^*_{i+1} such that all but α_{i+1} points of D_i are endvertices of some path and no such path has vertices in $\bigcup_{j=0}^{i} D_i$. (Let us recall that H^*_{i+1} differs from the subgraph of G induced by $\bigcup_{j=1}^{i+1} D_j$ only in D_{i+1}.) Now we can add the edges in S to the paths having common endvertices with them. As we know, all the endvertices of the considered paths are in D_i and all but α_{i+1} points of D_i are endvertices of at least one of them. By adding two edges of $S\subset H^*_{i+1}$ to a path P we obtain a circuit of H^*_{i+1} or a path which is embeddable in H^*_{i+1}. These latter paths we complete to circuits of H^*_{i+1}. $S\subset F$, the contracted graph F/S is a forest which corresponds to a forest on D_{i-1} in the natural way. (The image of a vertex in D_i is its neighbour in S.) The contracted graph is indeed a forest. We set $H_{i-1}=(\langle\bigcup_{j=0}^{i-1} D_j\rangle\setminus\langle D_{i-1}\rangle)\cup(F/S)$ which completes the general step.

In the last step we simply complete the paths of H_2 into circuits. In fact there are at most α_2 uncovered edges in V_1.

3.3. *Calculation*

The circuits of H^*_i can be extended to circuits of H_i, H_{i+1} and finally to circuits of G. Thus we obtain circuits of G which cover all edges having a covered image in some H^*_i.

In the i-th step we constructed

$$\sum_{j=1}^{f}\left[\frac{|V(K'_j)|}{2}\right]=\frac{|D_i|+|D_{i-1}|}{2}-\frac{|I_i|}{2}-\frac{\alpha_i}{2}-\frac{\beta_i}{2}$$

circuits and embeddeble paths such that:

(1) in $(G_i\cup E)\setminus F$ there are at most β_i uncovered edges.

(2) in $S\subset F$ there are at most α_{i+1} edges, uncovered by the circuits in H^*_{i+1} which are completions of embeddable paths.

The total number of constructed circuits is

$$\left(\frac{|D_r|}{2}-\frac{\alpha_{r+1}}{2}\right)+\sum_{i=2}^{r}\left(\frac{|D_i|+|D_{i-1}|}{2}-\frac{|I_i|}{2}-\frac{\alpha_i}{2}-\frac{\beta_i}{2}\right)$$

$$=n-1-\sum_{i=2}^{r}\left(\frac{|I_i|+\alpha_i+\beta_i}{2}\right)-\frac{\alpha_{r+1}}{2}-\frac{|D_1|}{2}=(n-1)-x\quad\text{(this defines } x\text{)}.$$

The number of uncovered edges of G is at most the sum of the numbers of uncovered edges in the H^*_i-s and the numbers of edges in the H_i-s having no images in the H^*_i-s, and the number of edges in the final forest $E\subset H_1$ (on D_1). This number is $\sum_{i=2}^{r}(|I_i|+\alpha_i+\beta_i)+\alpha_{r+1}+|D_1|-1$. (We have $|E|\le|D_1|-1$ for E is a forest.) G is 2-connected therefore any two edges of G can be covered by a circuit of G. We put the uncovered edges pairwise into circuits (if their number is odd then the last one remains alone). We obtained at most x such circuits which completes the calculation and the proof of Theorem 1.

4. The proof of Theorem 2

To prove this consequence of Theorem 1 we are going to use the theory of matroids. For definitions see [9].

Let $\{C_1, \dots, C_k\}$ be a covering of G by edges and circuits with k minimal.

Proposition 4.1. *The union of any m of the circuits and edges C_i covers at least $m+1$ vertices of G.*

Proof. $C_1, \dots, C_k$ is a minimal covering and we might apply Theorem 1 for the union of the m C_i-s. ∎

We define the matroid $C(\tilde{G})$ in the following way: If $e \in E(G)$ is contained by $C_{i_1}, \dots, C_{i_t}$ then we replace this edge by t other edges, namely $e_{i_1}, \dots, e_{i_t}$. That way we obtain a graph $\tilde{G}$ with multiple edges and $C(\tilde{G})$ is defined as its circuit matroid. Let $C_1^*, \dots, C_k^*$ denote the circuits and edges of G which correspond to the $C_1, \dots, C_k$ (i.e. $E(C_i^*) = \{e_j | e \in C_j\}$).

Proposition 4.2. *The union of m of the circuits and edges C_i^*, say H, has rank at least m in $C(\tilde{G})$.*

Proof. Let us denote the connected components of H by $H_1, \dots, H_s$. The C_i^* are connected so the H_j-s partition the m circuits and edges into s classes with say $m_1, \dots$ $\dots, m_s$ elements. It follows from Proposition 4.1 that H_j contains a tree of m_j edges consequently H contains a forest of m edges. We apply the following formula due to Nash-Williams (9): If $(S, \mathcal{M})$ is a matroid with rank function r, φ a homomorphism of $(S, \mathcal{M})$ to the matroid $(\varphi(S), \mathcal{M}\varphi)$ with rank function r_φ then

$$r_\varphi(X) = \min_{Y \subseteq X} \left(r(\varphi^{-1}(Y)) + |X - Y|\right) \quad \text{for} \quad X \subset \varphi(S). \quad ∎$$

4.3. The proof of Theorem 2

We apply the above formula to $C(\tilde{G})$ and $\varphi\colon C(\tilde{G}) \to \{1, \dots, k\}$ where $\varphi(e) = i$ iff $e \in C_i^*$. φ is well defined for the C_i^* form a disjoint covering of $\tilde{G}$. By Proposition 4.2 we have $r(\varphi^{-1}(X)) \geqq |X|$ for $X \subseteq \{1, \dots, k\}$, which gives us $r_\varphi(\varphi(C(\tilde{G}))) = |\varphi(C(\tilde{G}))| = k$. Therefore $\varphi(C(\tilde{G}))$ is an independent set so it is the image of an independent set of $C(\tilde{G})$ i.e. the image of a forest $\tilde{F} \subset \tilde{G}$. F corresponds to a forest $F \subset G$ having k edges. By the definition of $\tilde{G}$ this forest has the properties required in Theorem 2.

Problem. *Is Theorem 2 true for infinite graphs (or for a restricted class of infinite graphs)?*

5. The proof of Theorem 3

We have a graph G with a fixed distance tree W. Let us consider the graph G_i $(1 \leqq i \leqq r)$. $W_i \subset G_i$ therefore we might contract the edges of the star system W_i. We identify $V(G_i/W_i)$ with D_{i-1} in the natural way. Thus we obtain a map φ_i: $E(G_i \setminus W_i) \to E(G_i/W_i)$, $\varphi_i(e)$ is the image of $e \in E(G_i \setminus W_i)$ after the contraction of W_i. $\varphi_i(e)$ is always an edge because W_i consists of disjoint stars with their centers in D_{i-1} and D_i induces the empty subgraph in G_i.

Lemma 5.1. *Suppose that G has no circuit of length four, then for* $i=1, \ldots, r$

(a) If $\varphi_i(x, y)=\varphi_i(z, v)$ *then the edges* (x, y) *and* (z, v) *form a triangle in* G_i *with some* $e \in E(W_i)$.

(b) For any $e \in E(W_i)$ *we have at most one such triangle.*

Proof. (a) We have two possibilities

(1) both (x, y) and (z, v) are edges between a point of D_i and a point of D_{i-1} (let x, $v \in D_i$ and y, $z \in D_{i-1}$), or

(2) for one of the edges say (x, y) we have x, $y \in D_{i-1}$ $(v \in D_i,\ z \in D_{i-1})$. $x, y, z, v \in D_{i-1}$ is impossible for the contraction is identical on the edges of $\langle D_{i-1} \rangle \subset G_i$.

Case (1). If $y \neq z$ then $\varphi_i(x, y)=\varphi_i(z, v)=(y, z) \in E(G_i/W_i)$ is trivial which makes $x=v$ impossible. Consequently (x, z), $(y, v) \in E(W_i)$, these have to be contracted edges. But then (x, y, v, z) would be a $C_4 \subset G_i \subset G$, a contradiction. If $y=z$ then $\varphi_i(x, y)=(y, w) \in E(G_i/W_i)$ for some $w \in D_{i-1}$ and indeed (x, w), $(v, w) \in E(W_i)$. Therefore (y, x, w, v) would be a $C_4 \subset G_i \subset G$. Case (1) leads to contradiction.

Case (2). We have $\varphi_i(x, y)=\varphi_i(z, v)=(x, y) \in E(G_i/W_i)$ which means that $z \in \{x, y\}$, say $z=x$. Here $(v, y) \in W_i$, it has to be a contracted edge and (x, y, v) a triangle. This proves (a).

(b) Two triangles containing $e \in EW_i$ would imply the existence of a C_4 in G. ∎

Notation. If there exists a triangle (x, y, v) with $(y, v)=e \in W_i$ (using the above notation) then we refer to the unique (x, y) and (x, v) as e' and e''.

Proof of Theorem 3. Applying Theorem 4 (see Introduction) we can cover the graphs (G_i/W_i) $(2 \leqq i \leqq r)$ and G_{r+1} by $[|D_{i-1}|/2]$ and $[|D_r|/2]$ paths and circuits respectively. It follows that we can extend the circuits in G_i/W_i into circuits of G_i such that the contracted images of the latter ones are the circuits chosen in G_i/W_i and the paths in G_i/W_i into paths with their endvertices in D_{i-1}. By Lemma 5.1 the extended paths and circuits cover all the edges of G with the possible exception of W and $\{e', e'' | e \in W\}$. We also know that for $e \in W$ either e' or e'' is in an extended circuit or path H. If H does not contain e then it contains only one endvertex of e, say x, because the contracted image of H is a circuit (or path) if H is a circuit (or path). H contains one edge from the triangle (e, e', e''), exchanging this edge for the other two we obtain another circuit or path with the same contracted image as H. Consequently we might suppose that the extensions of the covering circuits and paths of the graphs G_i/W_i contain at least two edges of each triangle (e, e', e''). The endvertices of each extended path $P \subset G_i$ are in D_{i-1} and by Lemma 2.1e there is a circuit $C \subset G$ containing P. The extended circuits and the circuits containing the extended paths cover almost all edges of G. The uncovered edges are in a one-to-one correspondence with a subset of $E(W)$, where W is a tree with $|E(W)|=n-1$

$$\left[\frac{|D_1|}{2}\right]+\ldots+\left[\frac{|D_r|}{2}\right] \leqq \left[\frac{|D_1|+\ldots+|D_r|}{2}\right] = \left[\frac{n-1}{2}\right]$$

is the number of our circuits. The theorem follows.

Remark. Indeed Theorem 1 is a trivial consequence of Theorem 3 for C_4-free graphs.

Acknowledgement. The author wishes to thank L. Lovász for his continuous help.

References

[1] J. A. Bondy and U. S. R. Murty, *Graph theory with applications*, Problem 6.
[2] A. Donald, An upper bound for the path number of a graph, *Journal of Graph Theory*, **4** (1980), 189—201
[3] P. Erdős, A. W. Goodman and L. Pósa, The representation of a graph by set intersections, *Canad. J. Math.*, **18** (1966), 106—112.
[4] L. Lovász, On covering of graphs, in: *Theory of graphs, Proceedings of the Colloquium held at Tihany, Hungary, 1966*, (ed: P. Erdős and G.O.H. Katona), Academic Press, New York, 1968, 231—236.
[5] L. Lovász, A brief survey of matroid theory, *Matematikai Lapok*, **22** (1971), 249—267.
[6] L. Lovász, *Combinatorial problems and exercises*, North-Holland, 1979, 266—267, 332.
[7] C. St. J. A. Nash-Williams, An application of matroids to graph theory, *I. C. C. Rome/Dunod—Gordon and Breach*, 1967, 263—265.
[8] Tao Mao-qi, Shen Yun-qiu and Ku Tung-hsin, An Erdős conjecture—The planar case, *preprint*.
[9] D. J. A. Welsh, *Matroid theory*, Academic Press, London, New York, San Francisco, 1976.

L. Pyber

Mathematical Institute of the
Hungarian Academy of Sciences
Budapest, P.O.B. 127
1364, Hungary