COMBINATORICA 1 (1981) 13—24

MINIMAL DECOMPOSITIONS OF GRAPHS INTO MUTUALLY ISOMORPHIC SUBGRAPHS

by

F. R. K. CHUNG

Bell Laboratories
Murray Hill, New Jersey 07974, USA

P. ERDŐS*

Mathematical Institute of the Hungarian
Academy of Sciences
Budapest, Hungary H-1053
and

R. L. GRAHAM

Bell Laboratories
Murray Hill, New Jersey 07974, USA

Received 14 May 1979

Suppose $\mathcal{G}_n=\{G_1, \dots, G_k\}$ is a collection of graphs, all having n vertices and e edges. By a U-decomposition of $\mathcal{G}_n$ we mean a set of partitions of the edge sets $E(G_i)$ of the G_i, say $E(G_i)=\sum_{j=1}^{r} E_{ij}$, such that for each j, all the E_{ij}, $1\leqq i\leqq k$, are isomorphic as graphs. Define the function $U(\mathcal{G}_n)$ to be the least possible value of r a U-decomposition of $\mathcal{G}_n$ can have. Finally, let $U_k(n)$ denote the largest possible value $U(\mathcal{G})$ can assume where $\mathcal{G}$ ranges over all sets of k graphs having n vertices and the same (unspecified) number of edges.

In an earlier paper, the authors showed that

$$U_2(n) = \frac{2}{3}n+o(n).$$

In this paper, the value of $U_k(n)$ is investigated for $k>2$. It turns out rather unexpectedly that the leading term of $U_k(n)$ does not depend on k. In particular we show

$$U_k(n) = \frac{3}{4}n+o_k(n), \quad k \geqq 3.$$

1. Introduction

Let $\mathcal{G}=\{G_1, G_2, \dots, G_k\}$ be a collection of graphs,** all having the same number of edges. By a *U-decomposition of* $\mathcal{G}$ we mean a set of partitions of the edge sets $E(G_i)$ of the G_i, say $E(G_i)=\sum_{j=1}^{r} E_{ij}$, such that for each j, all the E_{ij}, $1\leqq i\leqq k$, are isomorphic as graphs. Under the above hypothesis, $\mathcal{G}$ always has such a decomposition since we can always take all the E_{ij} to be single edges. Define the function $U(\mathcal{G})$ to be the least possible value of r a U-decomposition of $\mathcal{G}$ can have. Finally, let $U_k(n)$ denote the largest possible value $U(\mathcal{G})$ can assume

* This work was performed while the author was a consultant at Bell Laboratories.

** In general we follow the terminology of [1].

AMS subject classification (1980): 05 C 35

where $\mathcal{G}$ ranges over all sets of k graphs each having n vertices and the same (unspecified) number of edges.

In previous work [2], [3], the function $U_2(n)$ was investigated rather thoroughly. In particular, it was shown that

$$U_2(n) = \frac{2}{3}n + o(n). \tag{1}$$

An example of a pair of graphs (G_1, G_2) achieving the bound in (1) is given by taking G_1 to be a star S_n with n edges and G_2 to be $\left(\frac{n}{3}\right)K_3$ (see Fig. 1).

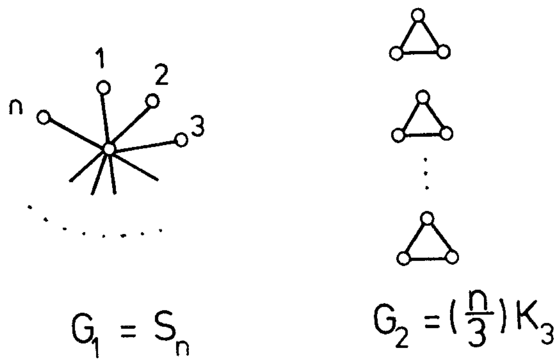

Fig. 1. A pair of graphs $\mathcal{G}$ with $U(\mathcal{G}) \sim \frac{2}{3}n$

We should point out here that strictly speaking, S_n has $n+1$ vertices and furthermore, when $n \not\equiv 0 \pmod 3$, $\left(\frac{n}{3}\right)K_3$ is undefined. However, here as throughout the entire paper, such statements are always to be taken with the understanding that the graphs may have to be adjusted slightly by adding or deleting (asymptotically) trivial subgraphs so as to make the stated assertion technically correct.

It was also shown in [2] that when $\mathcal{G}$ is restricted to *bipartite* graphs, the corresponding function $U_2^*(n)$ satisfies

$$U_2^*(n) = \frac{1}{2}n + o(n)$$

with an extreme pair given by taking G_1 to be a star $S_{\frac{n}{2}}$ (together with $\frac{n}{2}$ isolated vertices) and G_2 to be $\frac{n}{2}$ disjoint edges.

In this paper we study $U_k(n)$ for $k \geqq 3$. Already for the case $k=3$ it is not hard to find graphs G_1, G_2, G_3 on n vertices which require asymptotically more than $\frac{2}{3}n$ subgraphs in any U-decomposition. For example, taking $G_1 = S_n$, $G_2 = \left(\frac{n}{3}\right)K_3$

and $G_3=\left(\frac{n-\sqrt{n}}{2}\right)S_1\cup K_{\sqrt{n}}$, then for $\mathcal{G}=(G_1, G_2, G_3)$ we have

$$U(\mathcal{G}) = \frac{3}{4}n+o(n).$$

In fact, as we will show, this is the worst possible behavior since

$$U_3(n) = \frac{3}{4}n+o(n).$$

What was completely unexpected is that it does not get any worse than this as k increases. Indeed, the main result of the paper is that for any fixed $k\geqq 3$,

$$U_k(n) = \frac{3}{4}n+o(n).$$

Before proceeding to the proof of this, we remind the reader of the following notation: S_n denotes the star on n edges, i.e., the graph consisting of n vertices of degree 1 and one vertex of degree n; K_n denotes the complete graph on n vertices; nG denotes n disjoint copies of G; $G\subseteq H$ indicates that G is a (partial) subgraph of H; and finally $V(G)$ and $E(G)$ denote the vertex set and edge set, respectively, of a graph G and $v(G)$ and $e(G)$ denote the corresponding cardinalities $|V(G)|$ and $|E(G)|$.

2. The main result

The bulk of the paper will be devoted to proving the following result.

Theorem. *For any fixed* $k\geqq 3$,

$$U_k(n) = \frac{3}{4}n+o(n). \tag{3}$$

The proof of (3) is somewhat complicated. An outline of the plan of attack is as follows. We first choose an arbitrary fixed $\varepsilon>0$. We assume we begin with graphs $(G_1, \dots, G_k)=\mathcal{G}$ each having n vertices for a (sufficiently) large value of n, and e_0 edges. We will then successively remove isomorphic subgraphs H from the G_i, thereby decreasing the number e of edges currently remaining in each of the original graphs. Just what the subgraphs $H=H(e)$ are which will be removed will depend on the current value of e. There will be basically *six* distinct ranges for e, which we show in Fig. 2.

The STEP k notation indicates the process by which $H(e)$ is chosen. Each of the steps requires rather different arguments; the preparation for these arguments will now be made in a series of lemmas.

Let us denote by $h(\mathcal{G})$ the maximum number of edges in any subgraph H with $H\subseteq G_i$, $1\leqq i\leqq k$.

Lemma 1.

$$h(\mathcal{G}) \geqq \frac{e_0^k}{\binom{n}{2}^{k-1}}. \tag{4}$$

Proof. Let Λ_i denote the set of all 1—1 mappings of $V(G_i)$ into $V(G_1)$. For $\lambda_i \in \Lambda_i$, $e_i \in E(G_i)$, $1 \leqq i \leqq k$, define

$$I_{\lambda_2, \dots, \lambda_k}(e_1, \dots, e_k) = \begin{cases} 1 & \text{if } \lambda_i \text{ maps } e_i \text{ onto } e_1, \\ 0 & \text{otherwise,} \end{cases}$$

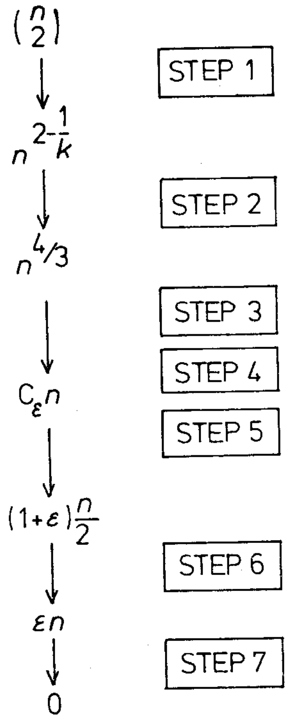

Fig. 2. Ranges for *e*

where we say that λ_i maps e_i onto e_1 if it maps the endpoints of e_i onto those of e_1. Then

$$S = \sum_{\substack{e_1 \in E(G_1) \\ \vdots \\ e_k \in E(G_k)}} \sum_{\substack{\lambda_2 \in \Lambda_2 \\ \vdots \\ \lambda_k \in \Lambda_k}} I_{\lambda_2, \dots, \lambda_k}(e_1, \dots, e_k) = \sum_{\substack{e_1 \in E(G_1) \\ \vdots \\ e_k \in E(G_k)}} (2(n-2)!)^{k-1} = e_0^k (2(n-2)!)^{k-1}.$$

Since $|\Lambda_i| = n!$ for all i, then for some $\bar{\lambda}_2 \in \Lambda_2, \ldots, \bar{\lambda}_k \in \Lambda_k$,

$$\sum_{\substack{e_1 \in E(G_1) \\ \vdots \\ e_k \in E(G_k)}} I_{\bar{\lambda}_2, \ldots, \bar{\lambda}_k}(e_1, \ldots, e_k) \geqq \frac{1}{|\Lambda_2|} \cdots \frac{1}{|\Lambda_k|} S \geqq$$

$$\geqq \frac{e_0^k}{(n!)^{k-1}} (2(n-2)!)^{k-1} = \frac{e_0^k}{\binom{n}{2}^{k-1}}.$$

The $\bar{\lambda}_i$ now determine a subgraph H common to all of the G_i which has at least $\dfrac{e_0^k}{\binom{n}{k}^{k-1}}$ edges and the lemma is proved. ▮

Suppose G is a graph with $v(G)=n$ and $e(G)=e=mn$.

Lemma 2.

(i) *If* $n^{1/3} \leqq m$ *then* $\left(\frac{1}{3}\sqrt{\frac{n}{m}}\right) S_m \subseteqq G$;

(ii) *If* $m < n^{1/3}$ *then* $\left(\frac{1}{3} m\right) S_m \subseteqq G$.

Proof. Let $X=\{x_1, \ldots, x_r\} \subseteqq V(G)$ denote the set of centers of a maximum set of disjoint S_m's. Let Y consist of all vertices in $V(G)-X$ which form the endpoints of these S_m's. Thus, $|Y|=rm$. Suppose $\deg(v) \geqq (r+1)(m+1)$ for some $v \in Y$. Since v is joined to at most r vertices of X and rm vertices of Y then v is connected to at least m vertices in $Z=V(G)-X-Y$, say $W=\{w_1, \ldots, w_m\}$. We now remove x_i, center of the S_m to which v belongs, and the m vertices in Y attached to x_i. Also, add v to X and W to Y. Thus, we still have $X-\{x_i\} \cup \{v\}$ as the centers of disjoint S_m's, the endpoints of which are in Y.

We now keep repeating this process. Suppose at some stage the vertex $x_i \in X$ we remove also has $\deg(x_i) \geqq (r+1)(m+1)$. Thus, x_i is connected to some vertex $u \in Z-W$ since

$$|X|-1+|Y|+|W| \leqq r-1+rm+m < \deg(x_i).$$

In this case we can add v to X, W to Y and also add u to Y (to give x_i a complete disjoint S_m of its own), forming $r+1$ disjoint S_m's in G. However, this contradicts the definition of r.

Thus, we must eventually reach a stage at which all $v \in Y$ have $\deg(v) < (r+1)(m+1)$. Since $|Z| \leqq n$ and $S_m \not\subseteqq Z$ then $e(Z) \leqq \frac{1}{2} mn$. The number of edges incident to some point in Y is at most

$$|Y| \cdot \max_{v \in Y} \deg(v) \leqq rm(r+1)(m+1).$$

Finally, the number of edges incident to some point in X is at most rn.

Therefore

$$e = mn \leqq rm(r+1)(m+1)+rn+e(Z)$$

$$\leqq rm(r+1)(m+1)+\frac{1}{2}mn+rn.$$

This implies

$$\frac{1}{2}mn \leqq r(r+1)m(m+1)+rn. \tag{5}$$

Case (i). $n^{1/3} \leqq m$. Suppose $r < \frac{1}{3}\sqrt{\frac{n}{m}}$. Then

$$r(r+1)m(m+1)+rn$$

$$\leqq \frac{1.1}{9}\frac{n}{m}\cdot m(m+1)+\frac{1}{3}\cdot\frac{n^{3/2}}{m^{1/2}}$$

$$\leqq n\left(\frac{0.4}{3}m+\frac{1}{3}m\right) < \frac{1}{2}mn$$

for n sufficiently large which contradicts (5).

Case (ii). $m < n^{1/3}$. Suppose $r < \frac{1}{3}m$. Then

$$r(r+1)m(m+1)+rn$$

$$\leqq \frac{1.1}{9}m^3(m+1)+\frac{1}{3}mn$$

$$\leqq m\left(\frac{1.2}{9}n+\frac{1}{3}n\right) < \frac{1}{2}mn$$

for n sufficiently large which again contradicts (5). This proves the lemma. ∎

Lemma 3. *If* $e(G) \geqq 2dt$ *then either*

$$S_d \subseteqq G \quad or \quad tS_1 \subseteqq G.$$

Proof. For $t=1$ the assertion is clear. Suppose for some $t>1$ that $S_d \not\subseteqq G$. Choose an edge $y \in E(G)$. Remove y and all incident edges from G, forming G'. Since $\deg(v) \leqq d-1$ for all $v \in V(G)$ then

$$e(G') \geqq 2dt-2d+1 > 2d(t-1).$$

By induction, $(t-1)S_1 \subseteqq G'$. Thus, since y is disjoint from this $(t-1)S_1$ then $tS_1 \subseteqq G$. ∎

Lemma 4. *If* $e(G) \geqq \frac{n}{2}+3dt$ *then either*

$$S_d \subseteqq G \quad or \quad tS_2 \subseteqq G.$$

Proof. For $t=1$ the assertion clearly holds. Suppose for some $t>1$ that $S_d \nsubseteq G$. Choose an $S_2 \subseteq G$. Remove it and all incident edges, forming G'. Since $\deg(v) \leq d-1$ for all $v \in V(G)$ then

$$e(G') \geq \frac{n}{2} + 3d(t-1).$$

Thus, by induction $(t-1)S_2 \subseteq G'$. Since the S_2 originally removed from G is disjoint from this $(t-1)S_2$ then $tS_2 \subseteq G$ and the lemma is proved. ∎

We are now ready for the proof of the Theorem. What we will do is to describe and analyze each step in the decomposition process as the current number of edges e passes through the previously indicated ranges. In particular, at any time $e \leq \varepsilon n$ we immediately go to STEP 7, which is simply the removal of subgraphs consisting of a single edge.

STEP 1: $n^{2-1/k} < e \leq \binom{n}{2}$. In this step, we repeatedly apply Lemma 1, removing a common subgraph having at least $\frac{e^k}{\binom{n}{2}^{k-1}}$ edges. Thus, if e_i denotes the number of edges remaining in each graph after i repetitions have been performed then

$$e_{i+1} \leq e_i - \frac{e_i^k}{\binom{n}{2}^{k-1}}. \tag{6}$$

Let $\alpha_i \equiv \frac{e_i}{\binom{n}{2}}$. Then $\alpha_0 = \frac{e}{\binom{n}{2}}$ and

$$\alpha_{i+1} \leq \alpha_i - \alpha_i^k \equiv f(\alpha_i).$$

Thus,

$$f'(x) = 1 - kx^{k-1}$$

and so, $f(x)$ achieves a maximum at $x_0 = \left(\frac{1}{k}\right)^{\frac{1}{k-1}}$ and $f(x)$ is monotone increasing for $0 \leq x < x_0$.

Suppose

$$\alpha_i \leq \left(\frac{1}{i}\right)^{\frac{1}{k-1}} \leq x_0 \quad \text{for some} \quad i \geq 1.$$

Then

$$f(\alpha_i) \leq f\left(\left(\frac{1}{i}\right)^{\frac{1}{k-1}}\right) = \left(\frac{1}{i}\right)^{\frac{1}{k-1}} - \left(\frac{1}{i}\right)^{\frac{k}{k-1}} = \left(\frac{1}{i}\right)^{\frac{1}{k-1}}\left(1-\frac{1}{i}\right) \leq \left(\frac{1}{i+1}\right)^{\frac{1}{k-1}}.$$

Therefore

$$\alpha_{i+1} \leq f(\alpha_i) \leq \left(\frac{1}{i+1}\right)^{\frac{1}{k-1}}.$$

Also, we have

$$\alpha_1 \leqq f(\alpha_0) \leqq f(x_0) = \left(\frac{1}{k}\right)^{\frac{1}{k-1}} \leqq 1$$

so that by induction,

$$\alpha_i \leqq \left(\frac{1}{i}\right)^{\frac{1}{k-1}} \quad \text{for all } i,$$

i.e.,

$$e_i \leqq \frac{\binom{n}{2}}{i^{\frac{1}{k-1}}}. \tag{7}$$

Choosing $i_0 = n^{1-1/k}$, we see that

$$e_{i_0} \leqq \binom{n}{2} \Big/ n^{\left(1-\frac{1}{k}\right)(k-1)^{-1}} \leqq n^{2-1/k}.$$

Thus, at most $n^{1-1/k}$ subgraphs are removed during STEP 1.

STEP 2: $n^{4/3} < e \leqq n^{2-1/k}$. In this step, we repeatedly apply Case (i) of Lemma 2. Abusing notation slightly, let e_0 denote the number edges each graph has at the beginning of this step. In general, if e_i denotes the number of edges remaining after i applications of the lemma, then

$$e_{i+1} \leqq e_i - \frac{1}{3}\sqrt{e_i} \tag{8}$$

since the number of edges in $\left(\frac{1}{3}\sqrt{\frac{n}{m}}\right) S_m$ is essentially $\frac{1}{3}\sqrt{nm} = \frac{1}{3}\sqrt{e}$. Let $e_i' = 9e_i$. Equation (8) then becomes

$$e_{i+1}' \leqq e_i' - \sqrt{e_i'} \equiv g(e_i'). \tag{9}$$

Note that $g(x) = x - \sqrt{x}$ is a parabola (at a 45° tilt) which is monotone increasing for $x \geqq 1/4$. Suppose for some $t \geqq i/2 > 0$ that

$$e_i' \leqq \left(t - \frac{i}{2}\right)^2.$$

Then

$$e_{i+1}' \leqq g(e_i') \leqq g\left(\left(t-\frac{1}{2}\right)^2\right) = \left(t-\frac{i}{2}\right)^2 - \left(t-\frac{i}{2}\right) = \left(t-\frac{i+1}{2}\right)^2 - \frac{1}{4} < \left(t-\frac{i+1}{2}\right)^2.$$

Since $e_0' \leqq n^{2-1/k}$ by hypothesis then taking $t = n^{1-1/2k}$ we have by induction

$$e_i < e_i' \leqq \left(n^{1-1/2k} - \frac{i}{2}\right)^2.$$

We apply this process only as long as $e_i > n^{4/3}$ so that at most $2(n^{1-1/2k} - n^{2/3})$ subgraphs are removed in this step.

STEP 3: $C_\varepsilon n < e \leqq n^{4/3}$ for a large constant C_ε depending on ε. In this step, we repeatedly apply Case (*ii*) of Lemma 2. Again, let e_i denote the number of edges remaining in each graph after Lemma 2 (*ii*) has been applied i times. Then

$$e_{i+1} \leqq e_i - \frac{1}{3}\left(\frac{e_i}{n}\right)^2. \tag{10}$$

By letting $\beta_i = e_i/3n^2$ the inequality becomes

$$\beta_{i+1} \leqq \beta_i - \beta_i^2.$$

By performing an analysis parallel to that used for the α_i in STEP 1 (with $k=2$), we deduce

$$\beta_i < \frac{3n^2}{i} \quad \text{for} \quad i \geqq 1. \tag{11}$$

Actually, we could take advantage of the fact that $\beta_0 \leqq n^{4/3}$ and strengthen (11) but it would have no effect on the final estimates. Hence, to reach $e \leqq C_\varepsilon n$ requires the removal of at most $\frac{3n}{C_\varepsilon}$ subgraphs.

STEP 4: $e \leqq C_\varepsilon n$. The first part of this step consists in successively removing $(\log n) S_1$ from all the graphs as long as possible. The second part consists of successively removing $S_{\log n}$ from all the graphs as long as possible. If after this process stops, the number e of remaining edges is less than εn then we go directly to STEP 7. Thus, we may assume that $e > \varepsilon n$. Let us denote by $H_1, \ldots, H_j$ those graphs having no $(\log n) S_1$ as a subgraph and by $H_{j+1}, \ldots, H_k$ those subgraphs having no $S_{\log n}$ as a subgraph. By Lemma 3, if $(\log n) S_1 \nsubseteq H_i$ then $S_{\frac{\varepsilon n}{2 \log n}} \subseteq H_i$, i.e., $S_{\log n} \subseteq H_i$ (with a similar argument applying if $S_{\log n} \nsubseteq H_i$). Since $e > \varepsilon n$ then we must have $1 \leqq j < k$. This completes STEP 4.

Note that the number of subgraphs removed in this step is at most $\frac{C_\varepsilon n}{\log n}$.

STEP 5: $\frac{n}{2}(1+\varepsilon) < e \leqq C_\varepsilon n$. Let Δ denote the largest degree of any vertex in any H_i, $1 \leqq i \leqq k$. Since $e > \frac{n}{2}(1+\varepsilon)$ then by Lemma 3, $(\log n) S_1 \nsubseteq H_1$ implies $S_{\frac{n}{4 \log n}} \subseteq H_1$, i.e.,

$$\Delta \geqq \frac{n}{4 \log n}. \tag{12}$$

Define

$X_i = \{v \in V(H_i)\colon\ \Delta - \deg(v) \leqq 1\}$,

Y_i = maximum set of disjoint S_2's in H_i.

By definition

$$e \geqq \frac{1}{2}|X_i|(\Delta - 1) \quad \text{for all } i,$$

i.e.,

$$|X_i| \leqq C'_\varepsilon \log n.$$

Also, for $j+1 \leqq i \leqq k$, since $S_{\log n} \nsubseteq H_i$ then by Lemma 4,

$$\left(\frac{\varepsilon}{6}\frac{n}{\log n}\right) S_2 \subseteqq H_i,$$

i.e.,

$$|Y_i| \geqq \frac{\varepsilon}{6}\frac{n}{\log n}, \quad j+1 \leqq i \leqq k. \tag{13}$$

Define $x^* = \max_{1 \leqq i \leqq j} |X_i|$. Thus

$$x^* \leqq C'_\varepsilon \log n.$$

For some $i_0 \leqq j$, $|X_{i_0}| = x^*$. Therefore

$$e = e(H_{i_0}) \geqq (\Delta-1)x^* - \binom{C'_\varepsilon \log n}{2}. \tag{14}$$

Now, define $Z_i = \{v \in V(H_i)\colon \deg(v) \geqq \sqrt{n}\}$ for $1 \leqq i \leqq j$. Suppose $|Z_i| \leqq x^*-1$. Consider the graph H'_i induced by $V(H_i) - Z_i$. Note that

$$(\log n) S_1 \nsubseteq H'_i$$

and

$$\begin{aligned} e(H'_i) &\geqq e(H_i) - |Z_i|\Delta \\ &\geqq (\Delta-1)x^* - \binom{C_\varepsilon \log n}{2} - (x^*-1)\Delta \\ &\geqq \Delta - x^* - C''_\varepsilon \log^2 n \\ &\geqq \frac{n}{4\log n} - C'''_\varepsilon \log^2 n. \end{aligned}$$

Thus, by Lemma 3, for

$$m = \left(\frac{n}{4\log n} - C'''_\varepsilon \log^2 n\right)\Big/ 2\log n \tag{15}$$

we have $S_m \subseteqq H'_i$. However, the expression in (15) exceeds $\sqrt{n}$ for large n which means that $S_{\sqrt{n}} \subseteqq H'_i$. This contradicts the definition of Z_i. Hence, we may assume

$$|Z_i| \geqq x^*. \tag{16}$$

Finally, for $1 \leqq i \leqq j$, we define X'_i to be $X_i \cup Z'_i$ where $Z'_i \subseteqq Z_i$ is disjoint from X_i and so that

$$|X'_i| = |X_i| + |Z'_i| = x^*$$

(this is always possible by (15)).

It is now easy to see that we can remove $x^* S_2$ from each H_i so that *Δ is decreased by 2*. This can be done by choosing each $x_i \in X'_i$ as a center for an S_2 for $1 \leqq i \leqq j$ (since $\deg(x_i) \geqq \sqrt{n}$ and $x^* \leqq C'_\varepsilon \log n$ then this is always possible). For $j+1 \leqq i \leqq k$, (13) guarantees that $x^* S_2 \subseteqq H_i$.

STEP 5 consists in successively removing $x^* S_2$'s (of course, each time the value of x^* may change) until $e \leqq \frac{n}{2}(1+\varepsilon)$. Each time a subgraph is removed, the maximum degree Δ decreases by 2.

STEP 6: $\varepsilon n \leqq e \leqq \frac{n}{2}(1+\varepsilon)$. The plan in this step is similar to that of the previous step. In this case we will remove each time the subgraph $x^* S_1$ so that Δ *always decreases by 1*. To see that this is possible, define X_i as in the previous step, i.e., $X_i = \{v \in V(H_i): \Delta\text{-deg}(v) \leqq 1\}$. For $j+1 \leqq i \leqq k$, define Y_i' to be a maximum set of disjoint S_1's in H_i. As before, it follows that

$$x^* = \max_{1 \leqq i \leqq j} |X_i| \leqq C_\varepsilon^* \log n$$

and

$$|Y_i'| \geqq \frac{\varepsilon}{2} \frac{n}{\log n}.$$

Defining Z_i as in the preceding step and extending X_i to X_i' with $|X_i'| = x^*$, $1 \leqq i \leqq j$, it is not hard to see that $x^* S_1$ can be removed from each H_i so that Δ decreases by 1.

STEP 6 consists in successively removing $x^* S_1$ in this way until $e \leqq \varepsilon n$.

STEP 7: $e < \varepsilon n$. This final step consists in successively removing S_1, the subgraph consisting of a single edge. Of course, in this step at most εn subgraphs are removed.

We are now ready to count the number N of subgraphs into which each of the original graphs has been partitioned. Let σ_i denote the number of subgraphs removed during STEP i. Then

$$\begin{aligned} N &\leqq \sum_{i=1}^{7} \sigma_i \\ &\leqq n^{1-1/k} + 2n^{1-1/2k} + \frac{3n}{C_\varepsilon} + \frac{C_\varepsilon n}{\log n} + \sigma_5 + \sigma_6 + \varepsilon n \leqq \\ &\leqq \sigma_5 + \sigma_6 + 2\varepsilon n \end{aligned}$$

for $C_\varepsilon > \frac{3}{\varepsilon}$ and n sufficiently large. However, because of the guaranteed reduction in Δ (which at the beginning of STEP 5 is certainly less than n), we have

$$2\sigma_5 + \sigma_6 < n. \tag{18}$$

Also, since in STEP 6 each subgraph has at least one edge and $\left(e \leqq \frac{n}{2}(1+\varepsilon)\right.$ during this process$\left.\right)$,

$$\sigma_6 \leqq \frac{n}{2}(1+\varepsilon). \tag{19}$$

Adding (18) and (19) we obtain

$$2(\sigma_5 + \sigma_6) < \frac{n}{2}(3+\varepsilon). \tag{20}$$

Substituting in (17), we have for $\mathcal{G}=(G_1, \ldots, G_k)$,

$$U(\mathcal{G}) \leqq N \leqq \left(\frac{3}{4}+3\varepsilon\right) n.$$

Since both ε and $\mathcal{G}$ were arbitrary then we conclude

$$U(n) \leqq \frac{3}{4} n+o(n).$$

Since we have already given an example of three graphs

$$\left(S_n, \binom{n}{3} K_3, \left(\frac{n-\sqrt{n}}{2}\right) S_1 \cup K_{\sqrt{n}}\right) = \mathcal{G} \quad \text{with} \quad U(\mathcal{G}) = \frac{3}{4} n + o(n)$$

then the Theorem follows. ▮

It would be interesting to know if the $o(n)$ term could be strengthened, say, to $O(1)$.

3. Concluding remarks

If we restrict all the G_i to be *bipartite* then it turns out that the bound on the corresponding function $H_k^*(n)$ is the same as that for $U_k(n)$ when $k \geqq 3$, in contrast to the bound we mentioned previously:

$$U_2^*(n) = \frac{n}{2}+o(n).$$

In other words,

$$U_k^*(n) = \frac{3}{4} n+o(n)$$

for all $k \geqq 3$. An example of three bipartite graphs which achieve this bound is given by taking $G_1=S_n$, $G_2=\left(\frac{n}{4}\right) K_{2,2}$ and $G_3=\left(\frac{n-\sqrt{2n}}{2}\right) S_1 \cup K_{\sqrt{n/2}, \sqrt{n/2}}$.

In another direction, one can ask the same questions for r-uniform hypergraphs. Here, the answers required are harder to obtain and are known with less precision. For example, in the case of two r-uniform hypergraphs on n vertices, say $\mathcal{H}=(H_1, H_2)$, it can be shown that r even,

$$c_1 n^{r/2} < U(\mathcal{H}) < c_2 n^{r/2}$$

for suitable positive constants c_i. This topic will be treated more fully in a later paper.

Finally, it is natural to ask how close Lemma 1 is to the "truth", i.e., is this essentially the right order for $h(G_1, \ldots, G_k)$? This too we leave for later.

References

[1] J. A. BONDY and U. S. R. MURTY, *Graph Theory with Applications,* American Elsevier, New York (1976).

[2] F. R. K. CHUNG, P. ERDŐS, R. L. GRAHAM, S. M. ULAM and F. F. YAO, Minimal decompositions of two graphs into pairwise isomorphic subgraphs, *Proc. 10th Southeastern Conf. on Comb., Graph Th. and Comp.*, (1979) 3—18.

[3] F. FRANCES YAO, Graph 2-isomorphism is *NP*-complete, *Inf. Proc. Letters* **9** (1979) 68—72.