On the decomposition of graphs into copies of $P_3 \cup tK_2$

Miri Priesler (Moreno)
&
Michael Tarsi

Computer Science Department
School of Mathematical Sciences
Tel-Aviv university
69978
Israel

Abstract. An H-decomposition of a graph G is a representation of G as an edge disjoint union of subgraphs, all of which are isomorphic to another graph H. We study the case where H is $P_3 \cup tK_2$ - the vertex disjoint union of a simple path of length 2 (edges) and t isolated edges - and prove that a set of three obviously necessary conditions for $G = (V, E)$ to admit an H-decomposition, is also sufficient if $|E|$ exceeds a certain function of t. A polynomial time algorithm to test H-decomposability of an input graph G immediately follows

1 Introduction

All graphs considered, are finite, undirected, with no loops and no multiple edges. Let G and H be two graphs. An H-decomposition of G is a representation of G as an edge disjoint union of subgraphs, all of which are isomorphic to H. Graph decomposition was, and still is, a rather popular research area. Intensive research has been done on many special cases, beginning with the classical work on "Steiner triple systems" (H is a triangle and G a complete graph) [11] and proceeding with hundreds of papers until present days. (see [2] for a partial list of references).

One of the most important results in the area is the following, obtained by R.M. Wilson [12]: For every graph $H = (V, E)$ and an integer $n \geq n_0(H)$, the complete graph K_n has an H-decomposition if and only if $|E|$ divides $\binom{n}{2}$ and $n-1$ is divisible by $g.c.d.\{d(x)|x \in V\}$.

Wilson's theorem implies that the existence of H-decomposition of the complete graph K_n is decidable for any fixed graph H, in $O(\log(n))$ time. Practically, even for the simple case where H is the complete graph on 6 vertices, there are still several dozens of integers n (all in the interval 100-1000), for which H-decomposability of K_n is yet undetermined. (The case where H is a small complete graph and $G = K_n$ was the first and most intensively studied). Wilson's theorem points out a family of graph decomposition problems, which can be solved in polynomial time. Another of the few results of that character is due to Y. Caro [5], which presented a polynomial time algorithm to solve decomposition problems where both G and H are trees.

We are interested here in the class of problems, each determined by a fixed graph H, where the other graph G is given as input.

Definition. *For a fixed graph H, the H-decomposition problem is stated as follows: Can an input graph G be represented as an edge disjoint union of subgraphs, all of which are isomorphic to H?*

I. Holyer [9] proved that H-decomposition is NPC for every complete graph $H = K_n$, $n \geq 3$. He also conjectured NP completeness whenever H consists of at least 3 edges. In its general form the assertion of that conjecture is false (assuming $P \neq NP$). Even before Holyer stated his conjecture, Brouwer and Wilson, in an unpublished paper [4] gave a polynomial time algorithm (although their result is not stated in that form) for the case where $H = tK_2$ (the union of t disjoint edges). The same result was obtained later, independently, by N. Alon [1], after the case $H = 3K_2$ had been studied by Bialostocki and Roditty [3]. Later, Favaron, Lonc and Truszczynski [7] have obtained a similar result for the case where $H = P_3 \cup K_2$ is the disjoint union of a single edge and a simple path of length 2 (we measure path length by edges). Holyer's conjecture might still hold if restricted to the case where H is connected, or equivalently, contains a connected component with at least 3 edges (this equivalence is proved in section 3 of [6]). In addition to his result about complete graphs, Holyer [8] also proved H-decomposition to be NPC whenever H is either a simple circuit or a simple path (at least 3 edges long). Recently Cohen and Tarsi [6] proved H-decomposition to be NPC for a family of graphs H, which contains all trees.

While discussing known results, mention should be made of a closely related topic, the H-factorization problem, where vertices, rather than edges, play the main role: Given an input graph $G = (V, E)$, is there a collection of vertex disjoint subgraphs of G, isomorphic to a fixed graph H, such that the union of their vertex sets is V? H-factorization is also known as generalized matching (a complete matching is indeed a K_2-factorization). The complexity status of this class of problems was studied by Kirkpatrick and Hell [10], who showed that H-factoriation is NPC, whenever H contains a connected component of at least three vertices and it is polynomial otherwise.

The partial results on H-decomposition seem to indicate the validity of a similar statement, which turns Holyer's conjecture into an 'if and only if' form:

Conjecture. *H-decomposition is NPC, henever H contains a connected component of at least three edges and it is polynomial otherwise.*

The subject of this research is the second half of that conjecture, that is, constructing polynomial time algorithms for H-decomposition, where each connected component of H is either a single edge or a simple path of length 2. Let such a graph with s components which are paths of length 2 and t isolated edges, usually referred to as $sP_3 \cup tK_2$, be shortly denoted by $D_{s,t}$.

As mentioned above, the existence of a polynomial time algorithm for $D_{s,t}$-decomposition is known for $(s,t) = (1,0)$ (trivial), $s = 0$ ([4], [1]) and $(s,t) =$

$(1,1)$ ([7]). In the following section we prove the existence of a polynomial time algorithm for $D_{1,t}$-decomposition, for any natural number t.

2 Necessary and asymptotically sufficient conditions for $D_{1,t}$-decomposition

Almost all (hundreds of) known results on graph decomposition are of the following scheme: A system of a few simple necessary conditions for a graph G to admit an H-decomposition are also sufficient if G is large enough. In most cases such statements are restricted to certain families of graphs G such as complete graphs, or complete multipartite graphs. A complete characterization of H-decomposability contains an explicit list of the exceptions, that is, those graphs G which satisfy the conditions, but admit no H-decomposition. Complete solutions of that kind, where G is not restricted to a certain family, are given for $H = D_{0,3}$ in [3] and for $H = D_{1,1}$ in [7].

Being interested here mainly in the complexity status of decomposition problems, there is no need to explicitly list the exceptions (the number of which might increase exponentially with the size of H). It suffices to show that a system of easily (polynomially) verified conditions, are also sufficient, if the size of G exceeds a certain value (dependent on H). Once such a system is found, it implies the existence of a polynomial time algorithm for H-decomposition. A result of this character is presented in [1] for $H = D_{0,t}$. We solve here the case where $H = D_{1,t}$ by proving the following:

Theorem 1. *The following necessary conditions for a graph $G = (V, E)$ to admit a $D_{1,t}$-decomposition, are also suficient if $|E| \geq f(t)$, where $f(t)$ is a certain function of t.*

1. $|E| = k(t+2)$ is an integer product of $t+2$ and
2. The number of connected components of G with an odd number of edges is at most kt and
3. Let $d_1, \ldots, d_{|V|}$ be the degree sequence of the vertices of G in decreasing order, then $\sum_{i=1}^{m} d_i \leq k(m+1)$, for each $m = 1, \ldots, |V|$.

Proof

Necessity: Necessity of 1 is obvious. Admitting a $D_{1,t}$-decomposition, E can be partitioned into k copies of $D_{1,t}$ each containing t isolated edges. An odd component of G contains at least one of these kt edges. Necessity of 2 immediately follows. Condition 3 is the summation of the degree sequence over k copies of $D_{1,t}$.

Sufficiency: The case $t = 1$ is proved in [7] . Our condition 3 is replaced there by the equivalent: $\Delta(G) \leq 2k$ with the exclusion of the existence of one edge adjacent to all other edges. Accordingly, we assume throughout this proof $t \geq 2$.

Our proof has two main phases. First we prove decomposability of E into k subgraphs, some of which are isomorphic to $D_{1,t}$, while the others are copies of $D_{0,t+2}$. In the second phase we show that the number of $D_{0,t+2}$ copies in such a decomposition can be reduced to 0.

Definitions and Notation for the Proof: Let $G = (V, E)$ be a graph on $|E| = k(t+2)$ edges. A $D_{0,t+2}, D_{1,t}$*-decomposition* of G is a partition $P = \{D_1, \ldots, D_k\}$ of E into k subgraphs, each isomorphic either to $D_{0,t+2}$ or to $D_{1,t}$. The subgraphs $D_1, \ldots, D_k$ are called *P-graphs*. The P-graphs isomorphic to $D_{0,t+2}$ are referred to as *matchings*, while the copies of $D_{1,t}$ are called *clusters*. The 2-edges component of a cluster $C \in P$ is called the *hook* of C. The other t edges of a cluster, as well as all $t+2$ edges of each matching, are referred to as *daggers*. The *drift* $dr(P)$ of a $D_{0,t+2}$,$D_{1,t}$-decomposition P is the number of matchings in P. Clearly a $D_{1,t}$-decomposition is a $D_{0,t+2}, D_{1,t}$-decomposition P with $dr(P) = 0$.

Two edges, or two subgraphs, which share at least one vertex, as well as two vertices which are endvertices of an edge, are said to be *adjacent* to each other. An edge and each of its endvertices are *incident* to each other.

A connected component of a (sub)graph G' consisting of a single edge is referred to as a *trivial* component of G'.

Proposition 1. *Let $G = (V, E)$ be a graph which satisfies conditions 1 and 3 of theorem 1, with $k \geq 8/3(t+2)-2$, then G admits a $D_{0,t+2}, D_{1,t}$-decomposition.*

Proof of Proposition 1: Let the *excess* $c(x)$ of a vertex $x \in V$ be defined as $d(x) - k$ if $d(x) > k$, or 0 otherwise ($d(x)$ stands for the degree of x). The *total excess* $c(G)$ is the sum of $c(x)$ over V. Select now for each vertex x with $d(x) > 0$ a set of $c(x)$ *excess edges* incident to x, such that no edge is selected as an excess edge of more than one vertex. This can be done for one vertex after another. According to condition 3, $c(G) \leq k$, hence the degree of each vertex x with $c(x) > 0$ is more than the total excess and $c(x)$ edges, not previously selected, can be taken out of those incident to x. Define now a new graph G', on the same edge set E, obtained from G by 'disconnecting' all the excess edges of every vertex $x \in V$ from x and inserting instead a new vertex y, as a common endvertex for all the 'disconnected' edges. The degree in G' of the original vertices of G is at most k and $d(y) = c(G) \leq k$. Thus, G' is of maximum degree $\Delta(G') \leq k$. Theorem 1 of [1] states that a graph F with $kt \geq 8/3(t^2) - 2t$ edges and maximum degree $\Delta(F) \leq k$, always admits a $D_{0,t}$-decomposition. Consequently, our graph G' admits a $D_{0,t+2}$-decomposition. That is, E can be partitioned into k $D_{0,t+2}$ subgraphs of G'.Since all excess edges of G share a common endvertex y of G', each of them belongs to a distinct set of that partition. In particular, two excess edges of the same vertex x of G belong to distinct sets. The degrees of vertices x with $d(x) > k$ in G are reduced to exactly k in G' and hence the k edges incident to each of these vertices in G' belong to all k sets of the partition. It turns out that, as subgraphs of G, each set of that partition P is either isomorphic to $D_{0,t+2}$ or isomorphic to $D_{1,t}$. In other words, P is a $D_{0,t+2}$, $D_{1,t}$-decomposition of G. ∎

In what follows we refer to a graph G_o which satisfies conditions 1,2,3 of theorem 1 and admits a $D_{0,t+2}$, $D_{1,t}$-decomposition. We also assume $|E| \geq f(t)$, where the bound $f(t)$ is explicitly determined later during the proof. Let P_0 be a $D_{0,t+2}$, $D_{1,t}$-decomposition of G_0 whose drift $dr(P_0)$ is minimum over all such decompositions. To complete the proof of theorem 1 we should show $dr(P_0) = 0$. The first step towards achieving this goal is:

Proposition 2. *The union of any two matchings of P_0 consists of $2(t+2)$ trivial components.*

Proof: Each component of the union of two matchings A, $B \in P_0$ is either a simple path or an even circuit and the edges along each component are alternating between A and B. If any component has at least four edges, then four consecutive edges belonging to A, B, A, B, in that order, can be exchanged between A and B to the new order A, A, B, B. That way both A and B are transformed into clusters. If there exists a component of size 3, that is a path of the pattern A, B, A, it can be transformed into A, A, B. Two distinct 2-edges components both of the form A, B can be changed to A, A and B, B and finally, if there exists only one 2-edges component, of the pattern A, B, it can be transformed into A, A, while another trivial component of A is moved into B. In all cases a new partition is obtained with lower drift than that of P_0 - a contradiction.

Our main tool for the reduction of $dr(P_0)$ is:

Proposition 3. *If $\Theta \subseteq P_0$ is a set of P_0-graphs, which contains a matching M and in addition to M, another matching, or at least two clusters, then the subgraph $U = \cup_{D \in \Theta} D$ has a connected component, which contains at least two daggers.*

Proof of Proposition 3: Assume on the contrary, that Θ is a maximal subset of P_0, which contradicts the proposition. If $\Theta = P_0$ then $U = G_0$ and thus no two daggers of P_0 belong to the same component. Due to the existence of at least one matching, it contradicts condition 2 of theorem 1. Hence, there exists a P_0-graph R in $P_0 - \Theta$. Since Θ is maximal, there existx a connected component T of $U \cup R$ which contains at least 2 daggers. Let (t_1, t_2) be a pair of daggers, such that the distance in T between t_1 and t_2 (the length of a shortest path which contains both) is minimum. Let D denote the set of all daggers of P-graphs in Θ. To simplify keeping track of the proof, we cut it at this point into several phases.

Claim 1. *Given any daggers $d_1, d_2, d_3 \in D$, we can assume that both d_1 and d_2 belong to the matching M. We can also assume that d_3 does not belong to M, unless M is the only matching in Θ and d_3 is adjacent to the hooks of all other P_0-graphs in Θ (Such a dagger, if it exists, is clearly unique).*

What we mean by "we can assume " is that if P_0 does not satisfy the above then the edges can be repartitioned among the P_0-graphs of Θ to obtain a new decomposition, with the same number of matchings and clusters, where d_1 and d_2, but not d_3, belong to the same matching M.

Proof of Claim 1: Let us take care first of d_1 and d_2. By proposition 2, all the matchings in Θ are pairwise vertex disjoint and hence, if either d_1 or d_2 (or both) belongs to a matching it can be exchanged with any edge of M. If d_1 belongs to a cluster C, then, as Θ is assumed to contradict proposition 3, at most one edge of M is adjacent to (the hook of) C. Since $|M| = t + 2 \geq 4$ there exists another edge $d \in M$, which is nonadjacent to C and it is not d_2. Exchanging d_1 and d between C and M and, if necessary, acting similarly on d_2, we obtain d_1, $d_2 \in M$ as required. Consider now d_3. Assuming $d_3 \in M$. If there exists a P_0-graph in $\Theta - \{M\}$, which is nonadjacent to d_3 then any dagger of that graph can be exchanged with d_3 to reach the required situation. If there is no such a P_0-graph in Θ then every $D \in \Theta - M$ is a cluster with its hook adjacent to d_3. This case is given in claim 1 as an exception. ∎

Claim 2. *We can assume that both t_1 and t_2 belong to M (this holds also if t_1 and t_2 are adjacent, in which case it turns M into a cluster in contradiction to $dr(P_0)$ being minimum and thus completes the proof of proposition 3 in that case).*

Proof of Claim 2: We can directly apply claim 1 to either t_1 or t_2 (or both) if it belongs to D. If $t_i \notin D$ then it is a dagger of R. If R is a matching then by proposition 2 it is vertex disjoint of M and thus we can reach $t_i \in M$ by repartitioning $R \cup M$.

It is left to take care of the case where R is a cluster and either t_1 or t_2, or, if things come to worst, both of them, are daggers of R: If three edges of M are incident each with one vertex of the hook of R then we can take the component which contains that hook to serve as the component T, which contains the daggers t_1,t_2, now obviously in M, as required. Thus, we can assume that at most two edges of M are adjacent to the hook of R. If no dagger of R is adjacent to an edge of M then there are at least two (4 minus 2 adjacent to the hook) edges of M nonadjacent to any edge of R. These two edges can be exchanged with t_1, t_2 between M and R (only one edge, if only one $t_i \in R$) to reach the requirement of claim 2.

Finally consider the case where there exists a dagger $r \in R$ adjacent to an edge $d_1 \in M$. The set of daggers D contains at least $2t + 4$ edges ($2t + 4$ in two matchings, or $3t + 2$, which by $t \geq 2$ is at least as many, in one matching and two clusters), each in a distinct connected component of U. There are only $2t + 3$ vertices in the cluster R. This implies that there exists a dagger $d_2 \in D$, belonging to a component of $U \cup R$, which contains no edge of R. By means of claim 1 we can assume that both d_1 and d_2 are in M. Exchanging r and d_2 between R and

M, there are now two adjacent edges r and d_1 in M. If there exists another edge $d_3 \in D$ adjacent to r then claim 1 allows us to assume $d_3 \notin M$ (in the case where d_3 is the exceptional edge mentioned in claim 1 we first switch between d_1 and d_3). Now r and d_1 form a P_3 component, which turns M into a cluster, in contradiction to $dr(P_0)$ being minimum. ∎

Now we are in a position to complete the proof of proposition 3, assuming $t_1, t_2 \in M$: If t_1 and t_2 are adjacent to each other (as a result of the transformation used in the proof of claim 2) then they form a hook for M and proposition 3 is proved (a contradiction to $dr(P_0)$ being minimum). Otherwise, let $\{t_1, a_1, \ldots, a_m, t_2\}$ be a shortest path containing t_1 and t_2. By the minimum condition on the distance between t_1 and t_2, none of $a_1, \ldots, a_m$ is a dagger and no dagger other than t_1, t_2 is adjacent to any of $a_1, \ldots, a_m$. In particular, a_1 is contained in the hook H of a cluster C. The subgraph $H \cup \{t_1\}$ can be repartitioned into a copy H' of P_3 and a single edge t_1', such that the distance between t_1' and t_2 is strictly smaller than between t_1 and t_2 (a separate case is where t_1, t_2 and a third edge $t_3 \in M$ each is incident to one vertex of H. in that case $H \cup \{t_1, t_2, t_3\}$ can be repartitioned into two hooks and a dagger, disjoint from one of them and thus transforming M into a cluster). The drift of the partition obtained as C is replaced by $(C - H) \cup H'$ and M by $(M - \{t_1\}) \cup \{t_1'\}$ is at most equal that of P_0. Iterating this process until the two daggers are adjacent to each other, yields a $D_{0,t+2}, D_{1,t}$-decomposition with drift $< dr(P_0)$, which completes the proof of proposition 3. ∎

Proposition 2 forces any two matchings of P_0 to be vertex disjoint, while proposition 3 prohibits the existence of two vertex disjoint matchings. The only escape from contradiction is thus:

Proposition 4.

$$dr(P_0) < 1$$

It remains to eliminate the possibility of a single matching in P_0.

Proposition 5. *Let $M, C_1, C_2 \in P_0$ be, in that order, a matching and two clusters, such that no edge of C_1 has its two endvertices both in M and no edge of C_2 has its two endvertices in $M \cup C_1$, then no two daggers of $\Theta = \{M, C_1, C_2\}$ belong to the same connected component of $U = M \cup C_1 \cup C_2$.*

Proof of Proposition 5: Consider first the matching M and the cluster C_1. If no edge of C_1 has both its endvertices in M, then each component of $M \cup C_1$ is a simple path, either 1,2 or 3 edges long, except for the hook of C_1, which can form a simple path, 2 to 6 edges long, or any component on 3 edges. Having this structure, $M \cup C_1$ can clearly be repartitioned into two clusters, in contradiction to $dr(P_0)$ being minimum, unless $M \cup C_1$ has exactly one nontrivial component, consisting of either 2 or 3 edges.

Consider now $U = M \cup C_1 \cup C_2$, where no edge of C_2 has its endvertices both in $M \cup C_1$. Each component of U is a simple path with 1,2 or 3 edges, except for the nontrivial component of $M \cup C_1$ and the hook of C_2, which might form either 1 or 2 larger components. A very simple case analysis shows that, again, a repartition of U into 3 clusters is possible, unless no component of U holds more than a single dagger. ∎

Combining propositions 3 and 5 we immediately obtain:

Proposition 6. *Let $M, C_1, C_2 \in P_0$ be, in that order, a matching and two clusters, then either an edge of C_1 belongs to the subgraph of G_0, induced by the vertices of M, or an edge of C_2 belongs to the subgraph induced by the vertices of $M \cup C_1$.*

Completing the proof of theorem 1: Let us assume, in contradiction to theorem 1, that $dr(P_0) > 0$. By proposition 4 this implies $dr(P_0) = 1$. Let M be the only matching in P_0. The subgraphs of G_0 induced by the $2t + 4$ vertices of M hold at most $\binom{2t+4}{2}$ edges. Hence, at most that many clusters of P_0 have an edge in this subgraph. Once $f(t)$ is selected to make k larger than that value, there exists a cluster $C_1 \in P_0$ with no edge in that subgraph. The same argument, where M is replaced by $M \cup C_1$, implies the existence of a cluster C_2 with no edge in the subgraph induced by $M \cup C_1$. However, this contradicts proposition 6 and thus the proof of theorem 1 is now completed. ∎

A remark about the bound $f(t)$: The bound $f(t)$ for the size of E, obtained from our proof above is about $8t^3$. A significantly lower value can be reached for the price of a much longer proof. In fact we managed, by means of a long case analysis to prove theorem 1 with $k \geq 8/3(t+2)-2$ (that is $f(t) = 8/3(t+2)^2-2(t+2)$). The bottle neck here is the bound of $|E| \geq 8/3t^2 - 2t$, given in [1] for $D_{0,t}$-decomposition, which we use in proposition 1 above. Any improvement of this bound can be used to obtain a lower value for $f(t)$. The smallest graph that we know of, which satisfies conditions 1,2,3 of theorem 1, but does not admit a $D_{1,t}$-decomposition is of size $k = 2t$.

References

1. N. Alon, *A note on the decomposition of graphs into isomorphic matchings*, Acta Math. Hung. **42** (1983), 221–223.
2. J.C. Bermond & D. Sotteau, *Graph decompositiom and G-designs*, Proc. of the 5th British combinatorial conference, Aberdeen (1975).
3. A. Bialostocki & Y. Roditty, *$3K_2$-decomposition of a graph*, Acta math. Acad. Sci. Hungar. **40** (1982), 201–208.
4. A. E. Brouwer & R. M. Wilson, "The decomposition of graphs into ladder graphs", Stiching mathematisch centrum (zn 97/80), Amsterdam, 1980.

5. Y. Caro, "Decomposition and Partition of trees into isomorphic subtrees", Tel-Aviv University (M.Sc Thesis), 1985.
6. E. Cohen & M. Tarsi, *NP-Completeness of Graph-Decomposition Problems*, Journal of Complexity 7 (1991), 200–212.
7. O. Favaron, Z. Lonc & M. Truszczynski, *Decomposition of graphs into graphs with three edges*, Ars Combinatoria **20** (1985), 125–146.
8. Ian Holyer, "The complexity of graph theory problems", (Ph.D dissertation), Churchill College, Cambridge, 1980.
9. Ian Holyer, *The NP-completeness of some edge partition problems*, Siam J. of comp. **10** (1981), 713–717.
10. D. G. Kirkpatrick & P. Hell, "On the completeness of a generalized matching problem", 10th annual ACM Symposium on Theory of Computing Machinery, New-York, 1978.
11. J. Steiner, *Combinatorische aufgabe*, Z. Reine Angew Math **45** (1853), 181–182.
12. R.M. Wilson, "Decomposition of a complete graph into subgraphs isomorphic to a given graph", Utilitas Math., Winnipeg, 1976.