

ON RESTRICTED TWO-FACTORS*

PAVOL HELL†‡, DAVID KIRKPATRICK†‡, JAN KRATOCHVÍL§, AND IGOR KRÍŽ§

Abstract. A two-factor of G consists of disjoint cycles that cover $V(G)$. The authors consider the existence problem for two-factors in which the cycles are restricted to having lengths from a prescribed (possibly infinite) set of integers. Theorems are presented which derive the existence of such restricted two-factors in G from their existence in $G - u$ and $G - v$. The possibility of such theorems is then related to the complexity of the corresponding existence problem. In particular, the only four cases in which polynomial algorithms can be expected (in the sense that all other cases are shown to be NP-hard) are identified.

Key words. cycles, two-factors, triangle-free, NP-complete, polynomial algorithm

AMS(MOS) subject classifications. 05C38, 05B40, 05C99

1. Introduction. A *two-factor* of a graph G is a subgraph F of G in which $V(F) = V(G)$ and every vertex has degree two. Evidently, a two-factor of G consists of disjoint cycles that cover $V(G)$. The existence question for two-factors has been successfully studied; there is an elegant criterion for deciding if a graph G admits a two-factor, [1], [17], as well as a polynomial algorithm to find such a two-factor (or determine that none exists) [6], [7].

We consider the existence question under the additional restriction that the lengths of all the cycles comprising the two-factor belong to a given set of integers L . Specifically, let $L \subseteq \{3, 4, \dots\}$ be a nonempty set, and let $L^- = \{3, 4, \dots\} - L$. An *L -restricted two-factor* of a graph G is a two-factor F of G with the property that each component of F is a cycle the length of which belongs to the set L . In the case $L^- = \emptyset$ all cycle lengths are permitted, and we refer to it as the case of *unrestricted two-factors*; this is the well-studied case as explained above. The case $L^- = \{3\}$ has also been studied [3] and is usually referred to as the case of *triangle-free two-factors*. Another important case we shall encounter is that of *square-free two-factors*, i.e., the case when $L^- = \{4\}$.

For unrestricted two-factors [8] reports the following fact (cf. also [13]): If $uv \in E(G)$ and if $G - u$ as well as $G - v$ have two-factors, then G also has a two-factor. (This follows from Tutte's f -factor theorem, cf. [8, 13]; a simpler direct proof is given in the next section.) We show that similar results also hold for L -restricted two-factors when $L \subseteq \{3, 4\}$, but for no other L .

Vornberger [16] showed that recognizing graphs that have an odd-restricted two-factor ($L = \{3, 5, 7, \dots\}$) is NP-complete. Other L -restricted two-factor problems were shown to be NP-complete in Cornuejols and Pulleyblank's paper [4] (some of these proofs being attributed to Papadimitriou). On the other hand, [3] reports a polynomial algorithm for finding a triangle-free two-factor (or for showing that none exists). As an application of our constructions we show that recognizing graphs that have an L -restricted two-factor is NP-hard unless $L^- \subseteq \{3, 4\}$, i.e., except for the four particular cases $L^- = \emptyset$, $L^- = \{3\}$, $L^- = \{4\}$, and $L^- = \{3, 4\}$. This subsumes all the previous NP-com-

* Received by the editors July 14, 1987; accepted for publication (in revised form) May 5, 1988.

† School of Computing Science, Simon Fraser University, Burnaby, British Columbia, Canada V5A 1S6.

‡ Department of Computer Science, University of British Columbia, Vancouver, British Columbia, Canada V6T 1W5.

§ Department of Applied Mathematics, Charles University, Prague, Czechoslovakia.

¶ The work of these authors was supported in part by the Natural Sciences and Engineering Research Council of Canada under grants A5075 and A3583.

pleteness results, and in view of the above-mentioned polynomial algorithms for the first two cases ($L^- = \emptyset$, $L^- = \{3\}$), it suggests that the remaining two cases ($L^- = \{4\}$, $L^- = \{3, 4\}$) may also admit such algorithms. As further evidence we mention in the last section some augmenting path theorems that can be proved for these four cases.

2. Positive results. Our objective in this section is to extend the result cited in the Introduction, restated as Theorem 1 below. We begin by giving a simple proof of it, which will be helpful in understanding the extensions. The original proof in [8], as well as that in [13], depends on Tutte's f -factor theorem; the proof given here is elementary.

In the following text we often refer to set operations and terminology (such as union, intersection, symmetric difference, or membership) applied to subgraphs; it is to be understood that they apply to the edge-sets of those subgraphs.

THEOREM 1. *If $uv \in E(G)$ and if both $G - u$ and $G - v$ have two-factors, then G also has a two-factor.*

Proof. Let F_u and F_v be two-factors of $G - u$ and $G - v$, respectively, and let H be their symmetric difference, $H = F_u \oplus F_v$. In H each vertex other than u or v is incident with an equal number of edges of F_u and of F_v (0, 1, or 2 of each). Let P be a maximal trail in H of the form $u = x_0, x_1, \dots, x_k$ with all $x_{2i}x_{2i+1} \in F_v$ and all $x_{2i-1}x_{2i} \in F_u$ ($i = 0, 1, \dots$). (Recall that a trail may revisit a vertex but not an edge.) By the maximality of P and the above observation on incidencies in F_u and F_v , it follows that $x_k = u$ or $x_k = v$. If $x_k = u$ then $F_u \oplus P$ is a two-factor of G ; if $x_k = v$ then $(F_u \oplus P) \cup \{uv\}$ is such a two-factor. \square

In subsequent proofs we shall also be using trails of the above type. If H is any graph containing $F_u \oplus F_v$, a trail $P: u = x_0, x_1, \dots, x_k$ in H , with all $x_{2i}x_{2i+1} \in F_v$ and all $x_{2i-1}x_{2i} \in F_u$ ($i = 0, 1, \dots$), shall be called an *alternating u -trail* in H . If P is such a trail, we shall denote by P_i ($i = 0, 1, \dots$) the alternating u -trail $u = x_0, x_1, \dots, x_i$.

THEOREM 2. *If $uv \in E(G)$ and if both $G - u$ and $G - v$ have triangle-free two-factors, then G also has a triangle-free two-factor.*

First we prove a slightly weaker version of Theorem 2. Let G^- denote the graph obtained from G by deleting the edge uv , and G^+ the graph obtained from G by subdividing the edge uv with one new vertex zero, i.e., by replacing the edge uv with the path $u0, 0v$ (for $0 \notin V(G)$). Thus G^- and G^+ are defined relative to a special edge uv ; however, uv will always be clear from the context.

THEOREM 3. *If $uv \in E(G)$ and if both $G - u$ and $G - v$ have triangle-free two-factors, then G^- or G^+ also has a triangle-free two-factor.*

Proof. Again let F_u and F_v be triangle-free two-factors of $G - u$ and $G - v$, respectively, and let $H = F_u \oplus F_v$. The idea here is very similar to the above proof. We shall again be considering alternating u -trails in H ; however, we must be careful to choose P in such a way that the new two-factor does not contain any triangles. Therefore let P be a maximal alternating u -trail $u = x_0, x_1, \dots, x_k$ in H satisfying, for each edge $x_{2i}x_{2i+1}$ ($i = 0, 1, \dots$), all of the following four conditions:

- (i) There is no vertex y with $yx_{2i} \in F_u - (F_v \cup P_{2i})$ and $yx_{2i+1} \in F_u - P_{2i}$;
- (ii) There is no vertex y with $yx_{2i} \in F_u - (F_v \cup P_{2i})$ and $yx_{2i+1} \in F_v \cap P_{2i}$;
- (iii) There is no vertex $y \neq x_{2i+2}$ with $yx_{2i} \in F_v \cap P_{2i}$ and $yx_{2i+1} \in F_u - (F_v \cup P_{2i})$;
- (iv) There is no vertex $y \neq x_{2i+2}$ with $yx_{2i} \in F_u \cap F_v$ and $yx_{2i+1} \in F_u - (F_v \cup P_{2i})$.

In Fig. 1, \in and \notin denote, respectively, the membership and nonmembership in P_{2i} (except for the obvious relation $x_{2i}x_{2i+1} \notin P_{2i}$, which we did not show in the figures). The edges of $F_u - F_v$ are wavy, and those of $F_v - F_u$ are straight. Edges of $F_u \cap F_v$ are both wavy and straight; edges of F_u (which could also belong to F_v or not) are wavy with an interrupted straight line.

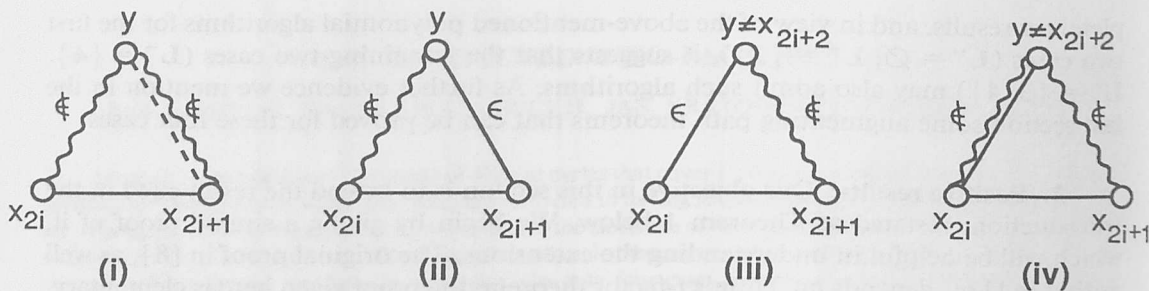


FIG. 1. The four forbidden configurations.

CLAIM 1. The trail P ends at $x_k = u$ or $x_k = v$.

Suppose $x_k \neq u, v$. Then x_k has an equal number (zero, one, or two) of incident edges of H in F_u and F_v ; thus P could not be maximal unless $k = 2i$ and each choice of $x_{2i}x_{2i+1} \in F_v - (F_u \cup P)$ violates one of (i)–(iv). Specifically, if $k = 2i - 1$ and P satisfies (i)–(iv), then letting $x_{2i-1}x_{2i}$ be any edge of $F_u - (F_v \cup P)$ extends P so that it still satisfies (i)–(iv), contrary to the maximality of P . Now, if $k = 2i$ and a particular choice of $x_{2i}x_{2i+1} \in F_v - (F_u \cup P)$ violates (i) or (ii), then there is a vertex y as described in the conditions and depicted in the figures. On the other hand, if some $x_{2i}x_{2i+1} \in F_v - (F_u \cup P)$ violates (iii) or (iv), then that means a violation occurs regardless of what choice of x_{2i+2} is made (in other words, in order to satisfy (iii), (iv), two distinct vertices would need to be x_{2i+2}). If there are two edges incident with x_{2i} in $F_v - (F_u \cup P)$, and if taking either as $x_{2i}x_{2i+1}$ violates one of (i)–(iv), then it could be because both violate one of (i), (ii), or because both violate one of (iii), (iv), or because one violates one of (i), (ii) and the other one of (iii), (iv). We shall treat the latter two cases together; taking into account also the possibility of there being only one edge incident with x_{2i} in $F_v - (F_u \cup P)$, we distinguish two situations:

(A) For each choice of $x_{2i}x_{2i+1}$ some vertex y violates (i) or (ii).

(B) For some choice of $x_{2i}x_{2i+1}$ two distinct vertices y, y^{\sim} are required to be x_{2i+2} by (iii) or (iv).

Clearly if neither (A) nor (B) occurs, then some choice of $x_{2i}x_{2i+1}$ leads to no violation of (i)–(ii) and at most one vertex y is required to be x_{2i+2} by (iii)–(iv). Therefore, letting x_{2i+2} equal such a vertex y if it exists, or terminating at x_{2i+1} if it does not exist, then the maximality of $P: u = x_0, x_1, \dots, x_{2i}$ is contradicted.

We now show that both (A) and (B) are impossible. In case (A) we have $x_{2i-1}x_{2i} \in F_u \cap P_{2i}$ while any violation of (i) or (ii) involves a vertex y with $yx_{2i} \in F_u - (F_v \cup P_{2i})$. Thus x_{2i} is incident in H with two edges of F_u and two edges of F_v . If the two edges of F_v are $x_{2i}z$ and $x_{2i}z^{\sim}$ then z or z^{\sim} are the choices for x_{2i+1} . Suppose $x_{2i+1} = z$ violates (i) or (ii) because of some y , and $x_{2i+1} = z^{\sim}$ violates (i) or (ii) because of some y^{\sim} . Clearly $y = y^{\sim}$ or else x_{2i} would be incident with three edges of F_u . For the same reason both $x_{2i+1} = z$ and $x_{2i+1} = z^{\sim}$ could not violate (i); thus we may assume that the choice $x_{2i+1} = z^{\sim}$ violates (ii). Then $yz^{\sim} \in F_v \cap P_{2i}$ yet no edge of F_u incident to y belongs to P_{2i} ; this is impossible as P_{2i} begins at $u \neq y$ and ends at $x_{2i} \neq y$. In case (B) we have $x_{2i}x_{2i+1} \in F_v$; if $y \neq y^{\sim}$ both violate (iii), (iv) then $x_{2i}y$ and $x_{2i}y^{\sim}$ are two more edges of F_v incident to x_{2i} , a contradiction. \square

CLAIM 2. If $x_k = u$, then $F_u \oplus P$ is a triangle-free two-factor of G^- . If $x_k = v$, then $(F_u \oplus P) \cup \{u0, 0v\}$ is a triangle-free two-factor of G^+ .

If $F_u \oplus P$ contains a triangle abc , then in G we have

(a) $ab \in F_v \cap P$ and $ac, bc \in F_u - P$, or

(b) $ab \in F_u - P$ and $ac, bc \in F_v \cap P$.

This follows from the definition of $F_u \oplus P$ and the fact that abc is not a triangle of F_u or of F_v . If $(F_u \oplus P) \cup \{u0, 0v\}$ contains a triangle abc , then abc does not contain the edges $u0, 0v$, and hence we also have (a) or (b) in G . Now we show that both (a) and (b) are impossible in G : In case (a) if, say, $a = x_{2i}$ and $b = x_{2i+1}$, then P violates (i) unless $ac \in F_u \cap F_v$ in which case $bc \in P$ according to (iv). In case (b) we may assume without loss of generality that bc precedes ac on P , i.e., that bc is an edge $x_j x_{j+1}$ and ac an edge $x_{j'} x_{j'+1}$ with $j < j'$. If $a = x_{2i}$ and $c = x_{2i+1}$, then P violates (ii), with $y = b$. If $c = x_{2i}$ and $a = x_{2i+1}$, then by (iii), $b = x_{2i+2}$, a contradiction to $ab \notin P$. \square

To prove Theorem 2 from Theorem 3, observe that G^- is a subgraph of G . Thus it only remains to consider the case when G^+ contains a triangle-free two-factor but G does not. In terms of the above proof, that means $(F_u \oplus P) \cup \{uv\}$ contains a triangle, i.e., that the path P constructed in that proof ends at $x_k = v$ and that $x_1 v \in F_u - P$. In this case G has an obvious triangle-free two-factor obtained from F_u by replacing $x_1 v$ with the path $x_1 u, uv$. \square

Note that Theorems 1 and 2 do not have a strict analogue for square-free two-factors, as can be seen by considering the graph K_4 with respect to any edge uv . (We thank Pierre Fraisse for this observation.) Nevertheless, Theorem 3 does generalize as follows.

THEOREM 4. *If $uv \in E(G)$ and if both $G - u$ and $G - v$ have square-free two-factors, then G^- or G^{++} also has a square-free two-factor.*

Here G^{++} is the graph obtained from G by subdividing the edge uv with two new vertices 0 and 1, i.e., by replacing uv with the path $u0, 01, 1v$ (for $0, 1 \notin V(G)$).

Proof. Again we assume that F_u and F_v are square-free two-factors of $G - u$ and $G - v$, respectively. However, instead of defining H as the symmetric difference, we let $H = F_u \cup F_v$; this will have the effect of allowing the edges of $F_u \cap F_v$ to be used in alternating trails. Specifically, let P be a maximal alternating u -trail in H , $P: u = x_0, x_1, \dots, x_k$, with each edge $x_{2i} x_{2i+1} \in F_v - F_u$ ($i = 0, 1, \dots$) and satisfying all of the following eight conditions:

- (v) There is no edge $yy' \in F_u - P_{2i}$ with $yx_{2i} \in F_u - (F_v \cup P_{2i})$ and $y'x_{2i+1} \in (F_v - F_u) \cap P_{2i}$;
- (vi) There is no edge $yy' \in (F_u \cap F_v) - P_{2i}$ with $yx_{2i} \in F_u - (F_v \cup P_{2i})$ and $y'x_{2i+1} \in F_v - (F_u \cup P_{2i})$;
- (vii) There is no edge $yy' \in (F_v - F_u) \cap P_{2i}$ with $yx_{2i} \in F_u - (F_v \cup P_{2i})$ and $y'x_{2i+1} \in F_v - F_u$;
- (viii) There is no edge $yy' \in F_u - P_{2i}$ with $y' \neq x_{2i+2}$ and $yx_{2i} \in F_u - P_{2i}$, $y'x_{2i+1} \in F_u - P_{2i}$;
- (ix) There is no edge $yy' \in (F_u \cap F_v) - P_{2i}$ with $y' \neq x_{2i+2}$ and $yx_{2i} \in F_v - (F_u \cup P_{2i})$, $y'x_{2i+1} \in F_u - (F_v \cup P_{2i})$;
- (x) There is no edge $yy' \in F_u - P_{2i}$ with $y' \neq x_{2i+2}$ and $yx_{2i} \in (F_v - F_u) \cap P_{2i}$, $y'x_{2i+1} \in F_u - (F_v \cup P_{2i})$;
- (xi) There is no edge $yy' \in (F_v - F_u) \cap P_{2i}$ with $y' \neq x_{2i+2}$ and $yx_{2i} \in F_u - P_{2i}$, $y'x_{2i+1} \in F_u - P_{2i}$;
- (xii) There is no edge $yy' \in (F_v - F_u) \cap P_{2i}$ with $y' \neq x_{2i+2}$ and $yx_{2i} \in F_v - F_u$, $y'x_{2i+1} \in F_u - (F_v \cup P_{2i})$.

In conditions (v)–(xii) it is to be assumed that yy' is disjoint from $x_{2i} x_{2i+1}$ as shown in Fig. 2. (We use the same conventions as in Fig. 1.)

So far we have encountered one basic difference from the previous proofs (in addition to an evident increase in complexity): the trail P can use the edges of $F_u \cap F_v$, although only in positions $x_{2i-1} x_{2i}$ ($i = 0, 1, \dots$). The second basic difference is the following: recall that the situation (B) from Claim 1 was shown impossible there. The analogous situation here can actually occur (see Fig. 3).

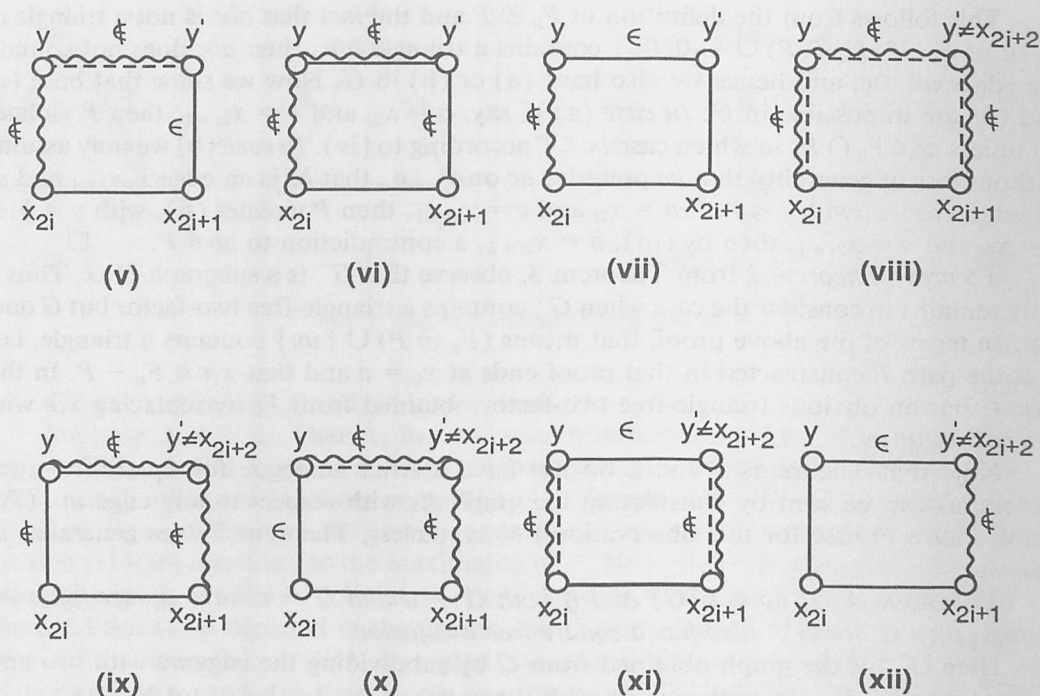
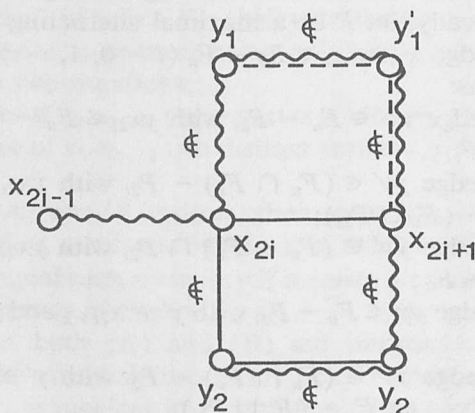


FIG. 2. The eight forbidden configurations.

FIG. 3. A choice of $x_{2i}x_{2i+1}$ that would require $x_{2i+2} = y'_1$ (by (viii)) and $x_{2i+2} = y'_2$ (by (ix)).

Nevertheless, the following claim remains true.

CLAIM 1. The trail P ends at $x_k = u$ or $x_k = v$.

As before, if $x_k \neq u, v$ then $k = 2i$, and for each choice of $x_{2i}x_{2i+1} \in F_v - (F_u \cup P)$ one of the following situations must occur:

(A) Some edge yy' violates (v), (vi), or (vii), or

(B) Some edges $y_1y'_1, y_2y'_2$ force two different choices $x_{2i+2} = y'_1$ and $x_{2i+2} = y'_2$ by (viii)–(xii).

We first prove that it is impossible that for all choices (there are at most two) of $x_{2i}x_{2i+1} \in F_v - (F_u \cup P)$ situation (A) occurs: assume that for $x_{2i}x_{2i+1}$ (v), (vi), or (vii)

is violated by yy' . Since $x_{2i-1}x_{2i} \in F_u \cap P_{2i}$, we have $y \neq x_{2i-1}$, and hence there is another choice $x_{2i}x_{2i+1} \in F_v - (F_u \cup P)$; assume that for $x_{2i}x_{2i+1}$ (v)–(vii) is violated by $y \sim y'$. Clearly $y \sim = y$, or else x_{2i} is incident with three edges of F_u . If $y \sim' \neq y'$, then for the same reason one of yy' and $yy \sim'$ is not in F_u ; hence, say, $x_{2i}x_{2i+1}$ violates (vii) and therefore $x_{2i}x_{2i+1}$ violates (v) or (vi). (Otherwise there is an unbalanced number of edges of $F_u \cap P$ and $F_v \cap P$ at $y = y \sim$.) This is not possible, as $y \neq u$, x_k would be incident with exactly one edge of P . If, on the other hand, $y' = y \sim'$ then $yy' \in F_u - F_v$ or else y' is incident with three edges of F_v . Thus both $y'x_{2i}$ and $y'x_{2i+1}$ are in $F_v \cap P_{2i}$ while $yy' \in P_{2i}$; this is impossible for $y' \neq u$, x_k .

As shown in Fig. 3, some choice of $x_{2i}x_{2i+1}$ may lead to (B). However, there is only a limited number of configurations that can cause this to happen: Assume that $x_{2i}x_{2i+1}$ leads to (B), i.e., that conditions (viii)–(xii) require two different choices $x_{2i+1} = y'_1$ and $x_{2i+1} = y'_2$ for some edges $y_1y'_1$, $y_2y'_2$. We first prove that $y_1 \neq y_2$ (in addition to $y'_1 \neq y'_2$). If $y_1 = y_2$ then

- $x_{2i}y_1 \in F_u - P_{2i}$, and we obtain a contradiction at y_1 (three edges of F_u , or an unbalanced number of edges of $F_u \cap P_{2i}$ and $F_v \cap P_{2i}$); or

- $x_{2i}y_1 \in F_v - P_{2i}$, and we have a similar contradiction (three edges of F_v incident to y_1); or

- $x_{2i}y_1 \in F_v \cap P_{2i}$, and we either note a similar contradiction again (three edges of F_v incident to y_1 , or an unbalanced number of edges of $F_u \cap P_{2i}$ and $F_v \cap P_{2i}$ at y_1), or, if both $y_1y'_1$ and $y_2y'_2$ contradict condition (x), we obtain the four-cycle $y_1y'_1x_{2i+1}y'_2y_1$, contradicting the fact that F_u is square-free.

Therefore $y_1 \neq y_2$, and since $x_{2i-1}x_{2i} \in F_u$, $x_{2i}x_{2i+1} \in F_v$, we may assume without loss of generality that $x_{2i}y_1 \in F_u$ and $x_{2i}y_2 \in F_v$ (as in Fig. 3). In fact, considering the position of x_{2i} on the path P_{2i-1} , and the fact that none of the edges $x_{2i}y$ in (viii)–(xii) belong to $F_u \cap P_{2i}$, we conclude that $x_{2i}y_1 \in F_u - P_{2i}$ and $x_{2i}y_2 \in F_v - P_{2i}$. This leaves the four possibilities obtained from combining each of (viii), (xi) with each of (ix), (xii); in addition to the case illustrated in Fig. 3, we list them in Fig. 4.

We have just proved the following subclaim.

SUBCLAIM. If $x_{2i}x_{2i+1}$ leads to (B) then H contains one of the four situations depicted in Figs. 3 and 4.

Note that in Figs. 3 and 4 we have marked the edge $x_{2i}y_2$ as $\notin P_{2i}$ even though (xii) does not require it; this is so because $x_{2i}y_2 \in P_{2i}$ would mean $x_{2i}y_2 \in P_{2i-1}$, contradicting the fact that x_{2i} is not incident with any edge of $F_u \cap P_{2i-1}$. (Recall that we are still assuming that $x_{2i} = x_k \neq u$, so that P_{2i-1} cannot begin at x_{2i} .)

In all four cases there is an obvious alternate choice, $x_{2i+1} = y_2$. We will be done if we can show that $x_{2i}x_{2i+1}$ does not lead to (A) or (B).

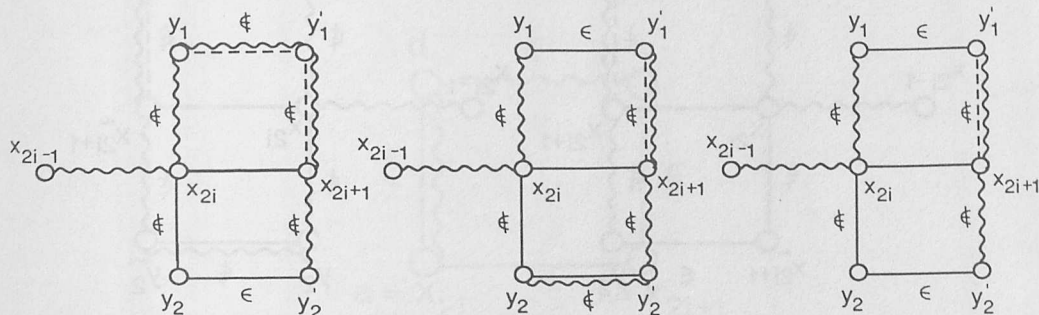


FIG. 4. The other configurations causing (B).

Case 1. $x_{2i}x_{2i+1}$ leads to (B).

According to the above subclaim, there is a configuration similar to one of the situations depicted in Figs. 3 and 4, on vertices $x_{2i-1}, x_{2i}, x_{2i+1}, y_1, y_1', y_2, y_2'$ (see Fig. 5).

By considering the degree of x_{2i} in F_u and F_v we conclude that $y_1 = y_1'$ and $y_2 = x_{2i+1}$. Note that it is not possible that both $y_1 y_1' \in F_u$ and $y_1 y_1' \in F_v$ because in such a case we would necessarily have $y_1' = y_1$ and $y_1 x_{2i+1} \in F_u, y_1' x_{2i+1} \in F_u$, yielding three edges of F_u at y_1 . Otherwise, say, $y_1 y_1' \in F_v \cap P$, as depicted in the example in Fig. 5, and, regardless of whether $y_1 y_1' \in F_u - P$ or $y_1 y_1' \in F_v \cap P$, we obtain a contradiction at $y_1 = y_1'$ arising from an unbalanced number of edges of $F_u \cap P$ and $F_v \cap P$. Thus Case 1 does not occur.

Case 2. $x_{2i}x_{2i+1}$ leads to (A).

Suppose $x_{2i}, x_{2i+1}, y_1, y_1'$ violate (v), (vi), or (vii). Then $y_1 = y_1'$ and $y_2 = y_2'$. If $y_1 y_1' \in F_u - P$ (as in (v) and (vi)) then we have a contradiction at y_1 because of the degree in F_u or an imbalance between $F_u \cap P$ and $F_v \cap P$. If $y_1 y_1' \in F_v \cap P$ (as in (vii)) we also have a contradiction because of an imbalance between $F_u \cap P$ and $F_v \cap P$ at y_1 . Thus Case 2 does not occur. This completes the proof of Claim 1. \square

CLAIM 2. If $x_k = u$, then $F_u \oplus P$ is a square-free two-factor of G^- . If $x_k = v$, then $(F_u \oplus P) \cup \{u0, 01, 1v\}$ is a square-free two-factor of G^{++} .

If $F_u \oplus P$ or $(F_u \oplus P) \cup \{u0, 01, 1v\}$ contains a four cycle $abcd$, then we may assume that in G we have the following:

- (a) $ab \in (F_v - F_u) \cap P, bc, cd, da \in F_u - P$; or
- (b) $ab, cd \in (F_v - F_u) \cap P, bc, da \in F_u - P$; or
- (c) $ab, bc \in (F_v - F_u) \cap P, cd, da \in F_u - P$; or
- (d) $ab, bc, cd \in (F_v - F_u) \cap P, da \in F_u - P$.

We now prove that each of these four situations is impossible.

(a) Without loss of generality $a = x_{2i}, b = x_{2i+1}$ for some i ; then $y = d, y' = c$ contradicts (viii) because $bc \notin P$ implies that $y' \neq x_{2i+2}$.

(b) Similarly, we may assume without loss of generality that cd precedes ab on P and that $a = x_{2i}, b = x_{2i+1}$ for some i (thus, $cd \in P_{2i}$); then, as in (a), $y = d, y' = c$ contradicts (xi).

(c) We may assume that, say, bc precedes ab on P ; but we must consider both possibilities, that $a = x_{2i}, b = x_{2i+1}$, or that $a = x_{2i+1}, b = x_{2i}$ (for some i). In the first

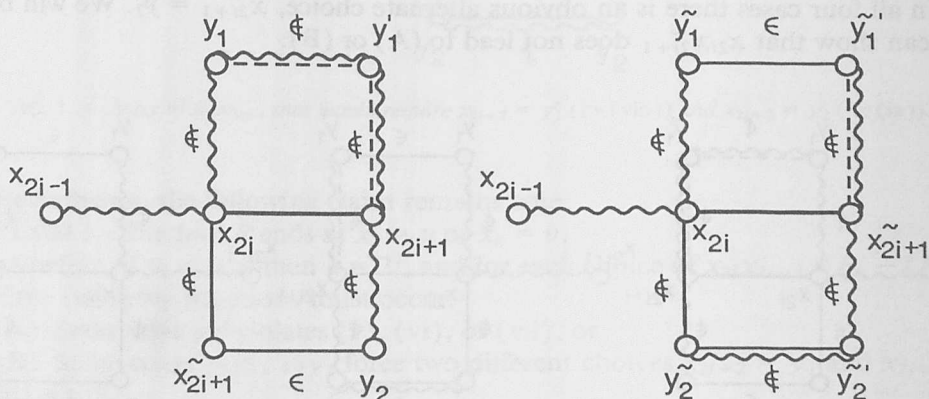


FIG. 5. Both $x_{2i}x_{2i+1}$ and $x_{2i}x_{2i+1}$ lead to (B). (This is only an example; there are 16 possible combinations to illustrate.)

case ($a = x_{2i}$, $b = x_{2i+1}$), $y = d$, $y' = c$ contradicts (v), unless $da \in F_u \cap F_v$ (see Fig. 6). In this situation (when $da \in F_u \cap F_v$) we must have $cd \notin F_v$, because F_v is square-free. Now $bc = x_{2j}x_{2j+1}$ (for some $j < i$): either $c = x_{2j}$, $b = x_{2j+1}$ and $y = d$, $y' = a$ contradicts (vi), or $b = x_{2j}$, $c = x_{2j+1}$ and $y = a$, $y' = d$ contradicts (ix). (For both of these contradictions we must remember to consider memberships in P_{2j} and not P_{2i} .) In the *second case* ($b = x_{2i}$, $a = x_{2i+1}$), $y = c$, $y' = d$ contradicts (x), unless $da \in F_u \cap F_v$, which is the situation dealt with above. Thus (c) is also impossible.

(d) We first prove that there is no four-cycle $prst$ in G such that

$$pr \in F_u - (F_v \cup P_{2i}), \quad rs, tp \in (F_v - F_u) \cap P_{2i}, \quad st \in F_v - (F_u \cup P_{2i})$$

for any i . Indeed, if such a four-cycle $prst$ were to exist in G , then we may assume that sr precedes pt on P and that $pt = x_{2j}x_{2j+1}$ for some $j < i$. If $p = x_{2j}$ and $t = x_{2j+1}$, then $y = r$, $y' = s$ contradict (vii); if $t = x_{2j}$ and $p = x_{2j+1}$ then $y = s$, $y' = r$ contradict (xii). Thus such a four-cycle $prst$ cannot exist. It now follows that if (d) were to occur in G , then bc cannot come after both ab and cd on P . (If $bc = x_{2i}x_{2i+1}$ and $ab, cd \in P_{2i}$ then $abcd = rstp$ yields a contradiction.) Hence we may assume without loss of generality that $ab = x_{2i}x_{2i+1}$ and $bc, cd \in P_{2i}$: if $a = x_{2i}$ and $b = x_{2i+1}$ then $y = d$, $y' = c$ contradict (vii); if $a = x_{2i+1}$ and $b = x_{2i}$ then $y = c$, $y' = d$ contradict (xii). This proves that (d) is impossible; Claim 2 and Theorem 4 follow. \square

Let G be a graph and $uv \in E(G)$. We denote by \mathcal{G}^* the set of all graphs obtained from G by subdividing the edge uv by any (finite) number of vertices. In the remainder of the paper we shall often refer to elements of \mathcal{G}^* . For simplicity we shall say *some* (any, all) G^* to mean *some* (any, all) $G^* \in \mathcal{G}^*$.

THEOREM 5. Let $L^- \subseteq \{3, 4\}$. If $uv \in G$ and if both $G - u$ and $G - v$ have L -restricted two-factors, then G^- or some G^* also has an L -restricted two-factor.

Proof. For the purposes of this proof, G^* may be chosen as follows: if L^- is nonempty and finite, the edge uv is subdivided with $\max(L^-) - 2$ (or more) new vertices; otherwise $G^* = G$.

When $L^- = \emptyset$, $G^* = G$ and Theorem 1 gives the desired conclusion. When $L^- = \{3\}$, $G^* = G^+$ and we apply Theorem 3. When $L^- = \{4\}$, $G^* = G^{++}$ and we conclude by Theorem 4. Finally, when $L^- = \{3, 4\}$, we take L -restricted two-factors F_u of $G - u$ and F_v of $G - v$, let $H = F_u \cup F_v$, and define $P: u = x_0, \dots, x_k$ to be a maximal alternating u -trail in H with each edge $x_{2i}x_{2i+1} \in F_v - F_u$ ($i = 0, 1, \dots$) and satisfying all 12 constraints (i)–(xii).

CLAIM 1. The trail ends at $x_k = u$ or $x_k = v$.

Otherwise, $k = 2i$ and for each choice of $x_{2i}x_{2i+1}$, we have the following:

(A) Some vertex y violates (i)–(ii), or some edge yy' violates (v)–(vii); or

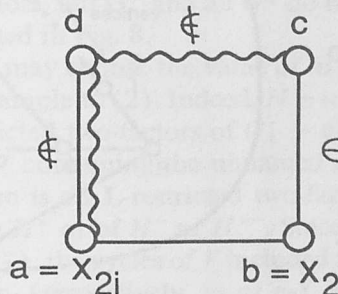


FIG. 6. A possible way to create (c). (The membership sign \in refers to membership in P_{2i} .)

(B) Some two vertices $y'_1 \neq y'_2$ are required to be x_{2i+2} by (iii)–(iv) or (viii)–(xii).

It again turns out that not every choice of $x_{2i}x_{2i+1}$ can lead to (A) or (B). Most of the work in verifying this is given in the preceding proof; it only needs to be checked that adding the rules (i)–(iv) does not change the situation. These verifications are similar and shall be omitted here. \square

CLAIM 2. If $x_k = u$, then $F_u \oplus P$ is an L -restricted two-factor of G^- . If $x_k = v$, then $(F_u \oplus P) \cup \{u0, 01, 1v\}$ is an L -restricted two-factor of G^* .

Since P satisfies both (i)–(iv) and (v)–(xii), Claim 2 follows from the corresponding claims in the proofs of Theorems 3 and 4. \square

3. Negative results. We begin by showing that Theorem 5 is the best possible in the following sense.

THEOREM 6. For any nonempty set $L \subseteq \{3, 4, \dots\}$ with $L^- \not\subseteq \{3, 4\}$ there exists a graph G with an edge uv such that both $G - u$ and $G - v$ have L -restricted two-factors, but neither G^- nor any G^* has an L -restricted two-factor.

Proof. We define, for all $\lambda \geq 3$ and all $p \geq 0, q \geq 0$, the graphs D_λ and $P_{p,q}$ (see Fig. 7).

Let L be finite and let $\lambda = \max L$. Then $G = D_\lambda$ satisfies the conditions in Theorem 6. Indeed, both $D_\lambda - u$ and $D_\lambda - v$ are just λ -cycles, while any two-factor of D_λ^- or of any D_λ^* is a cycle with more than λ vertices.

Let there exist nonnegative integers p and q such that $k = p + q + 5 \in L^-$ and $n = p + 2q + 9 \in L$. Then $G = P_{p,q}$ satisfies the conditions in Theorem 6. In fact, $(P_{p,q} - u) - \{ab, a'b'\}$ is a cycle of length n , and $(P_{p,q} - v) - \{tb', t'b\}$ is also a cycle of length n . On the other hand, any two-factor F of G^- or of any G^* must contain the paths $t \cdots ra$, and $t' \cdots r'a'$. In fact, F must also contain the path $b \cdots b'$; this is obvious when $p \neq 0$, and for $p = 0$ follows from the fact there is otherwise the forced cycle $t \cdots rabt' \cdots r'a'b't$, which misses u and v . Moreover, F must also contain at least one of the edges $tb', t'b$; otherwise F would miss v . If F contains both tb' and $t'b$, then F must miss u . So, without loss of generality, $tb' \in F$ and $t'b \notin F$. Then the cycle $t \cdots rab \cdots b't$ in F is of the forbidden length k . Since this cycle does not contain the special edge uv , the argument applies both to G^- and to any G^* , and hence they cannot have an L -restricted two-factor.

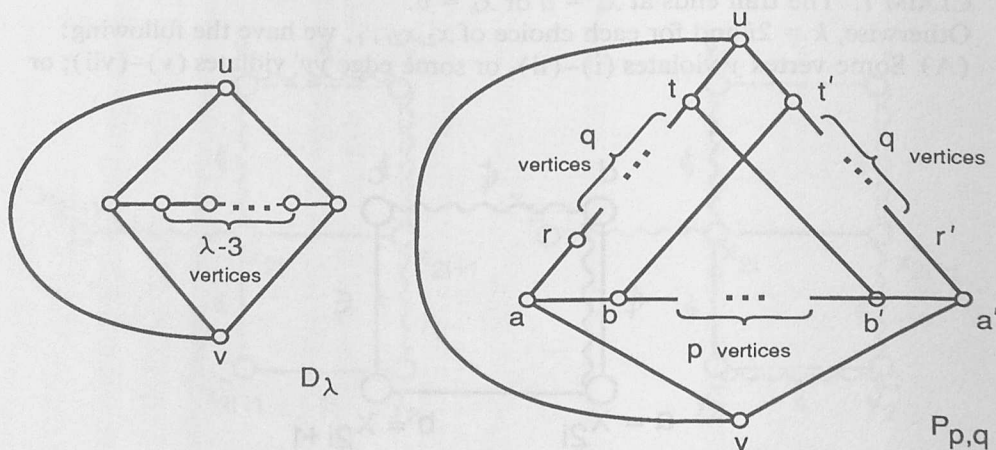


FIG. 7. The graphs D_λ and $P_{p,q}$.

When L^- is finite, $L^- \not\subseteq \{3, 4\}$, then letting $k = \max L^-$ and $n = k + 4$ implies that $k = p + q + 5$ and $n = p + 2q + 9$ for $q = 0$ and $p = k - 5$. Thus $G = P_{p,q}$ satisfies Theorem 6.

In the remaining case when both L and L^- are infinite, there must also always exist nonnegative integers p and q such that $p + q + 5 \in L^-$ and $p + 2q + 9 \in L$. Otherwise $k \in L^-$, $k \geq 5$, implies $k + 4 \in L^-$, $k + 5 \in L^-$, \dots , $2k - 1 \in L^-$. Therefore either some $k \in L^-$ has $k \geq 9$ and $\max L \leq k + 3$, or $\max L^- \leq 8$. \square

The fact that a graph G with a special edge uv satisfies Theorem 6 is closely related to a somewhat stronger property of G , which will turn out to be crucial to prove the apparent intractability of restricted two-factor problems. In addition to the notation $G^- = G - uv$, we also introduce here the notation $G^{--} = G - \{u, v\}$. We say that G is an L -clamp (with respect to $uv \in E(G)$), if $G - u$ and $G - v$ have L -restricted two-factors but G^- , G^{--} , and all G^* do not. It is easy to see that when L is finite and $\lambda = \max L$, the graph $G = D_\lambda$ is actually an L -clamp. Interestingly, it is also true that each graph $G = P_{p,q}$ with $p \geq 1$ is an L -clamp, provided $p + q + 5 \in L^-$ and $p + 2q + 9 \in L$. However, this fails when $p = 0$, because of the cycle $t \cdots rabt' \cdots r'a'b't$, which is an L -restricted two-factor of G^{--} . Nevertheless, the following general construction (Lemma 7) shows that the existence of a graph G satisfying Theorem 6 (i.e., a counterexample to Theorem 5) implies the existence of a clamp.

The graph H is said to be a *modular extension* of G^- if H contains G^- as an induced subgraph and no vertex of G^{--} is adjacent to a vertex of $H - G^-$. The condition that no G^* has an L -restricted two-factor is equivalent to the assertion that every L -restricted two-factor F of any modular extension H of G^- has each cycle entirely contained in G^- or in $H - G^{--}$. We say that G^- is L -coherent if neither G^- nor any G^* has an L -restricted two-factor, i.e., if in every L -restricted two-factor F of any modular extension H of G^- , the union of the cycles of F that are included in G^- misses at least one of the vertices u, v .

LEMMA 7. For any $L \subseteq \{3, 4, \dots\}$ the following three statements are equivalent:

(1) If $G - u$ and $G - v$ have L -restricted two-factors, then G^- or some G^* has an L -restricted two-factor.

(2) If $G - u$ and $G - v$ have L -restricted two-factors, then G^- or G^{--} or some G^* has an L -restricted two-factor.

(3) If $G - u$ and $G - v$ have L -restricted two-factors, then (both G^- and G^{--}) or (some G^*) has an L -restricted two-factor.

In other words, for any L either all (1)–(3) hold or none holds.

Proof. If L is finite then the graph D_λ (with $\lambda = \max L$) shows that all three statements are false; so we may assume that L is infinite.

Clearly (3) implies (1) and (1) implies (2). To prove that (2) implies (1), suppose that G is a counterexample to (1), i.e., a graph G with edge uv such that $G - u$ and $G - v$ have L -restricted two-factors, but G^- and all G^* do not. Thus G^- is L -coherent. Construct a graph H as illustrated in Fig. 8.

Since L is infinite, we may choose the value of $m \geq 1$ so that $m + 5 \in L$. We now claim that H is a counterexample to (2). Indeed, $H - u$ has an obvious L -restricted two-factor consisting of L -restricted two-factors of $G_1^- - v_1$ and $G_2^- - v_2$ together with the lower $(m + 5)$ -cycle of H^- containing the unnamed central vertices x, y , and v_1, v_2 , and v . Symmetrically, there is an L -restricted two-factor in $H - v$. Let F be any L -restricted two-factor of any H^* or of H^- or H^{--} . Since all of these graphs are modular extensions of G_1^- (and of G_2^-), the cycles of F included in G_1^- (respectively, in G_2^-) must miss at least one of u_1 or v_1 (respectively, u_2 or v_2). But, since the path $x \cdots y$ must belong to F , exactly one of the edges xu_1, xv_2 (exactly one of the edges yu_2, yv_1 , respec-

tively) belongs to F , and hence only one of u_1 and v_2 (respectively, only one of u_2 and v_1) can be missed by the cycles of F included in G_1^- (respectively, in G_2^-). Thus F contains either the cycle $xu_1uu_2y \cdots x$ or the cycle $xv_2vv_1y \cdots x$, missing either v or u , a contradiction.

To prove that (1) implies (3) we assume that $G - u$ and $G - v$ have \mathbf{L} -restricted two-factors but each G^* does not. Then we construct H again as in Fig. 8. Since both $H - u$ and $H - v$ have \mathbf{L} -restricted two-factors (cf. above), (1) implies that H^- or some H^* also has an \mathbf{L} -restricted two-factor. It is not hard to see that the former case is impossible, and that in the latter case both G^- and G^{--} must have \mathbf{L} -restricted two-factors. \square

The equivalence of (1) and (2) is crucial to our constructions. We remark on (3) only because it seems to come out of the same construction. In fact, exchanging the names u_2 and v_1 in H^- (cf. Fig. 8), and identifying as before, we obtain a new graph H' . Arguing in H' in the same vein as we did in H we conclude that (1) also implies the following:

(4) If both G^- and G^{--} have \mathbf{L} -restricted two-factors, then some G^* has an \mathbf{L} -restricted two-factor, or both $G - u$ and $G - v$ have \mathbf{L} -restricted two-factors.

This is worth restating for the case $\mathbf{L}^- \subseteq \{3, 4\}$, when we know (1) holds. Note that this includes the case of unrestricted two-factors.

COROLLARY 8. *If $\mathbf{L}^- \subseteq \{3, 4\}$, and if no G^* has an \mathbf{L} -restricted two-factor, then $G - u$ and $G - v$ have \mathbf{L} -restricted two-factors if and only if G^- and G^{--} have \mathbf{L} -restricted two-factors.* \square

We have now shown that if $\mathbf{L}^- \not\subseteq \{3, 4\}$, then there is an \mathbf{L} -clamp; we shall denote one such \mathbf{L} -clamp by G_L . We can now show our last construction (see Fig. 9).

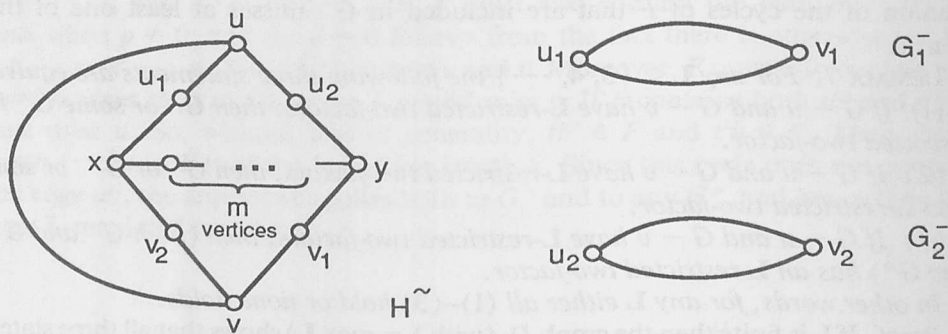


FIG. 8. The graph H is obtained from H^- and G_1^- , G_2^- by identifying all pairs of vertices with equal labels.

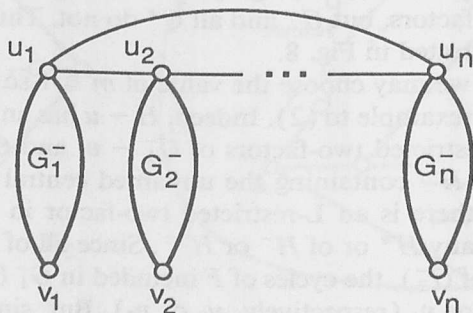


FIG. 9. The graph H_L .

The graph H_L consists of n copies of G_L^- , labeled G_i^- ($i = 1, 2, \dots, n$), with the vertices u_i joined in a cycle of length n . The value of n is chosen so that $n \in L$. Note that H_L is a modular extension of each G_i^- .

LEMMA 9. *Let $L^- \not\subseteq \{3, 4\}$, and let H_L be an induced subgraph of some graph H in which no vertex of $H_L - \{v_1, v_2, \dots, v_n\}$ is adjacent to any vertex of $H - H_L$. If F is an L -restricted two-factor of H then every cycle C of F containing a vertex of $H_L - \{v_1, v_2, \dots, v_n\}$ lies entirely inside H_L . Moreover, if one v_i lies in a cycle contained in H_L then all do.*

Proof. Since each G_i^- is L -coherent and H is a modular extension of G_i^- , the union of those cycles of F that are included in G_i^- contains at most one of the vertices u_i and v_i . Since G_i^- does not have an L -restricted two-factor, the union of these cycles must miss precisely one of u_i and v_i . Suppose v_i belongs to a cycle C of F that is included in H_L . Then by the coherence of G_i^- , C is included in G_i^- . Hence u_i is missed by all the cycles of F that are included in G_i^- , so the cycle $u_1 u_2 \dots u_n u_1$ must belong to F ; i.e., each u_j must be missed by the cycles of F belonging to G_j^- . Therefore all v_j must belong to cycles of F contained in H_L . \square

We also note that the graph H_L has an L -restricted two-factor: it is enough to combine the cycle $u_1 u_2 \dots u_n u_1$ with L -restricted two-factors of all $G_i^- - u_i$. Moreover, the graph $H_L - \{v_1, v_2, \dots, v_n\}$ also has an L -restricted two-factor consisting of the corresponding factors of all $G_i^- - v_i$. These properties, together with Lemma 9, will be sufficient to ensure that, when $L^- \not\subseteq \{3, 4\}$, the L -restricted two-factor problem is NP-hard.

Since L is a fixed set of lengths, determining if, say, the graph $G = C_n$ has an L -restricted two-factor amounts to testing the membership of n in L . Thus we cannot in general conclude that these problems are in NP, unless the description of L is easily given, e.g., for finite sets L . (In fact, some of these problems are undecidable [10]; cf. [12].) This is the reason we only aim to prove the problems are NP-hard. Notice that these results do translate to NP-completeness results for all the cases when testing membership in L is in NP (e.g., finite L or finite L^-). (An alternative approach, cf. [5], would be to consider the permitted cycles, viewed as subgraphs, part of the input; then all these problems are indeed in NP.)

THEOREM 10. *When $L^- \not\subseteq \{3, 4\}$ the L -restricted two-factor problem is NP-hard.*

Proof. Let $n \in L$. Since $n \geq 3$, the n -dimensional matching problem is NP-hard. In that problem we are given a disjoint union of n sets, $V = V_1 \cup V_2 \cup \dots \cup V_n$ and a collection E of subsets $e \subseteq V$ with the property that each $e \cap V_i$ has exactly one element; we are asked whether there is a subcollection $E' \subseteq E$ of the e 's that partition V . Given an instance of the n -dimensional matching problem we take a copy H_e of H_L for every $e \in E$, identifying vertex v_i in H_e with the unique element of $e \cap V_i$. It follows from the properties of H_L explained above that the resulting graph has an L -restricted two-factor if and only if the original problem has an n -dimensional matching (cf. [14], [15], or [9, p. 68], for similar arguments). \square

4. Conclusions and future directions. When $L^- \subseteq \{3, 4\}$ we have shown that L -restricted two-factors of $G - u$ and $G - v$ may be used to find an L -restricted two-factor of some G^* or of G^- (and G^{--}). In fact, the proofs we have given can be interpreted as linear-time algorithms for finding such factors. (The existence of such factors can be proved even for infinite graphs.) When $L^- \not\subseteq \{3, 4\}$ we have constructed examples G having L -restricted two-factors in $G - u$ and $G - v$, but none in any G^* or G^- (and G^{--}). These were then used to show that the corresponding L -restricted two-factor problems are NP-hard. Thus we can show the apparent intractability of all L -restricted two-factor problems, except for the following four cases:

- (1) $L^- = \emptyset$;
- (2) $L^- = \{3\}$;
- (3) $L^- = \{4\}$;
- (4) $L^- = \{3, 4\}$.

It is well known that the unrestricted two-factor problem (1) is in P , [7]. Recently, Cornuejols, Hartvigsen, and Pulleyblank [3] have shown that the triangle-free two-factor problem (2) is also in P . We take our positive results to be an indication that the remaining two L -restricted two-factor problems (3) and (4) should also admit polynomial algorithms. In a future paper we shall offer further evidence of this by proving an augmenting path theorem for each of the cases (1)–(4). In fact, these alternating “paths” are trails, very much like our alternating u -trails, satisfying the appropriate subset of (i)–(xii), except that the straight edges (here representing edges of a particular two-factor F_v of $G - v$) can be any edges of $G - F_u$. At this time we do not know how to find these augmenting trails in polynomial time.

REFERENCES

- [1] H. B. BELCK, *Reguläre Factoren von Graphen*, J. Reine Angew. Math., 188 (1950), pp. 228–252.
- [2] C. BERGE, *Sur le couplage maximum d'un graphe*, C.R. Acad. Sci. Paris Ser. 1. Math., 247 (1958), pp. 258–259.
- [3] G. CORNUEJOLS, D. HARTVIGSEN, AND W. PULLEYBLANK, *Triangle-free two-factors*, personal communication, 1986.
- [4] G. CORNUEJOLS AND W. PULLEYBLANK, *A matching problem with side conditions*, Discrete Math., 29 (1980), pp. 135–159.
- [5] G. CORNUEJOLS, D. HARTVIGSEN, AND W. PULLEYBLANK, *Packing subgraphs in a graph*, Oper. Res. Lett., 4 (1982), pp. 139–143.
- [6] J. EDMONDS, *Paths, trees, and flowers*, Canad. J. Math., 17 (1965), pp. 449–467.
- [7] J. EDMONDS AND E. L. JOHNSON, *Matching: a well-solved class of integer linear programs*, in Combinatorial Structures and Their Applications, R. Guy, H. Hanani, N. Sauer, and J. Schönheim, eds., Gordon and Breach, New York, 1970, pp. 89–92.
- [8] P. FRAISSE, P. HELL, AND D. KIRKPATRICK, *A note on f -factors in directed and undirected graphs*, Graphs and Combinatorics, 2 (1986), pp. 61–66.
- [9] M. GAREY AND D. S. JOHNSON, *Computers and Intractability*, W. H. Freeman, San Francisco, 1979.
- [10] J. E. HOPCROFT AND J. D. ULLMAN, *Introduction to Automata Theory, Languages, and Computation*, Addison-Wesley, Reading, MA, 1979.
- [11] P. HELL AND D. KIRKPATRICK, *Packing by cliques and by finite families of graphs*, Discrete Math., 49 (1984), pp. 118–133.
- [12] ———, *Packings by complete bipartite graphs*, SIAM J. Algebraic Discrete Methods, 7 (1986), pp. 199–209.
- [13] P. KATERINIS, *Some conditions for the existence of f -factors*, J. Graph Theory, 9 (1985), pp. 513–521.
- [14] D. KIRKPATRICK AND P. HELL, *On the completeness of a generalized matching problem*, in Proc. 10th Annual ACM Symposium on Theory of Computing, San Diego, CA, 1978, pp. 240–245.
- [15] ———, *On the complexity of general graph factor problems*, SIAM J. Comput., 12 (1983), pp. 601–609.
- [16] O. VORNBERGER, *Komplexität von Wegeproblemen in Graphen*, Reihe Theoretische Informatik, Universität Paderborn, 5 (1979).
- [17] W. TUTTE, *The factors of graphs*, Canad. J. Math., 4 (1952), pp. 314–328.