RICE UNIVERSITY

Restricted 2–factors in Bipartite Graphs

by

Summer Michele Husband

A THESIS SUBMITTED
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE

Master of Arts

APPROVED, THESIS COMMITTEE:

Nathaniel Dean, Associate Professor,
Chair

William Symes, Noah Harding Professor

Richard Tapia, Noah Harding Professor

6100 Main St. MS 134
Houston, TX 77005-1892

October, 2000

UMI Number: 1405675

Restricted 2–factors in Bipartite Graphs

Summer Michele Husband

Abstract

The k–restricted 2–factor problem is that of finding a spanning subgraph consisting of disjoint cycles with no cycle of length less than or equal to k. It is a generalization of the well known Hamilton cycle problem and is equivalent to this problem when $\frac{n}{2} \leq k \leq n-1$. This paper considers necessary and sufficient conditions, algorithms, and polyhedral conditions for 2–factors in bipartite graphs and restricted 2–factors in bipartite graphs. We introduce a generalization of the necessary and sufficient condition for 4–restricted 2–factors in bipartite graphs to $2k$–restricted 2–factors in bipartite graphs of a particular form.

Contents

Illustrations

1 Introduction

The *2-factor* problem is to find a spanning subgraph consisting of disjoint cycles. The *k-restricted 2-factor problem* is that of finding a 2-factor with no cycles of length k or less. Therefore, this problem is a generalization of the Hamilton cycle problem. However, the Hamilton cycle problem is NP-complete even for bipartite graphs. The k-restricted 2-factor problem in bipartite graphs is polynomial for $k = 2, 4$ and is open for $k \geq 6$. The weighted version of the k-restricted 2-factor problem is a generalization of the traveling salesman problem.

The problem of restricted 2-factors in bipartite graphs is interesting because it is similar to the well-known matching problem. Given a graph $G = (V, E)$, a *matching* is a set $M \subseteq E$ such that no two distinct edges in M share an endpoint. A *perfect matching*, or *1-factor*, is a matching M such that every node in V is incident to an edge in M (i.e., every node in (V, M) has degree 1). Other generalizations of the matching problem have had some of the same structure as the matching problem, so it is likely that restricted 2-factors in bipartite graphs will have similar structure. In fact, for 4-restricted 2-factors in bipartite graphs some "matching-type" theorems have already been proven (Hartvigsen [3], Király [4]).

Several open problems exist in the area of restricted 2-factors. Let P_k denote the problem of finding a 2-factor with no cycles of length k or smaller. Several complexity results are known for P_k in general graphs. Deciding whether a graph has a k-restricted 2-factor is known to be NP-complete for $k \geq 5$. A polynomial time algorithm is known for P_3, and P_4 is open. For bipartite graphs, P_4 is known to be polynomial, and P_k is open for $k \geq 6$ (Hartvigsen [3]). Given a graph with nonnegative weights, let ***weighted*** P_k denote the problem of finding a k–restricted 2–factor such that the sum of the weights of its edges is maximum. Vornberger showed that weighted P_4 is NP–hard (Hartvigsen [3]). Weighted P_3 is still an open problem. In addition, weighted P_k for bipartite graphs is open for all values of k.

Rosenfeld posed an interesting question concerning cubic, 3–connected planar graphs. It has been shown that every such graph has an even 2–factor (every cycle in the 2–factor is of even length). Rosenfeld asked whether these graphs have a 4–restricted even 2–factor (Rosenfeld [6]). To date, neither a proof nor a counterexample has been presented.

This paper summarizes results by Hartvigsen and Király on 2–factors in bipartite graphs. We will assume that all the given graphs are simple unless indicated otherwise. Note that loops are not possible since the graphs are bipartite. In the first section we consider a "Tutte–type" result for P_4 in bipartite graphs. This is followed by our generalization for P_{2k} in bipartite graphs of a particular form. Next, algorithms for determining 2-factors in bipartite graphs are defined. Finally, in the last section we consider the polyhedral structure of k–restricted 2–factors.

2 Tutte–type Theorems

In this section we consider necessary and sufficient conditions for 2–factors and restricted 2–factors in bipartite graphs. The theorems are similar in flavor to Tutte's well–known result on matchings, and they suggest an approach for P_{2k} in bipartite graphs. If $S \subseteq V$, then $G-S$ is the subgraph induced by $V \backslash S$. Let $oc(G-S)$ denote the number of components of $G-S$ with an odd number of nodes.

Theorem 1 *(Tutte's Theorem [5]) A graph G has a perfect matching if and only if $oc(G-Z) \leq |Z|$ for every $Z \subseteq V$.*

Tutte also has some fundamental results on n–factors and f–factors in general graphs. An *n-factor* of a graph G is a spanning subgraph in which every node has degree n. For a graph $G=(V,E)$, if H is a spanning subgraph of G, let $deg_H(v)$ denote the degree of $v \in V$ in H. Let $f : V \to \boldsymbol{Z}^+$ be a function from the nodes of G to the nonnegative integers. A subgraph H of G is called an *f-factor* if $deg_H(v) = f(v)$ for every $v \in V$. Hence, a 2–factor is an n–factor in which $n=2$ and an f–factor in which $f(v)=2$ for every node v. For a set $Z \subseteq V$, we define $f(Z) = \sum_{v \in Z} f(v)$. For $X, Y \subseteq V$, let $\delta(X,Y)$ denote the set of edges with one endpoint in X and the other in Y. Tutte proved the following two theorems.

Theorem 2 *(Tutte [7]) G has an n-factor if and only if for every two disjoint subsets X and Y of V, the number of those connected components K of $G-X-Y$ for which*

$n|V(K)| + |\delta(V(K),Y)|$ is odd does not exceed

$$n|X| + \sum_{v \in Y}(deg_G(v) - n) - |\delta(X,Y)|.$$

Theorem 3 is just a generalization of Theorem 2 to f–factors.

Theorem 3 *(Tutte [5]) A graph G has an f–factor if and only if for every two disjoint subsets X and Y of V, the number of those connected components K of $G - X - Y$ for which $f(V(K)) + |\delta(V(K),Y)|$ is odd, does not exceed*

$$f(X) + \sum_{v \in Y}(deg_G(v) - f(v)) - |\delta(X,Y)|.$$

For a graph $G = (V, E)$, let $A(G)$ denote the components of G that consist of a single node and let $B(G)$ denote the components of G that consist of two nodes joined by a single edge.

Theorem 4 *(Hartvigsen [3]) A bipartite graph $G = (V, E)$ has a 2–factor if and only if for every $Z \subseteq V$,*

$$|Z| \geq |A(G - Z)| + |B(G - Z)|.$$

Theorem 4 is essentially Tutte's n–factor theorem for $n = 2$ and restricted to the bipartite case. The necessity of the theorem is trivial to prove. The sufficiency follows from an Edmonds–type algorithm that is given in [3]. Note that although

Hartvigsen's theorem is for simple graphs, bipartite graphs cannot have loops. Also, if we let $B(G)$ denote the components of G with exactly two nodes, then Thereom 4 works for bipartite graphs with multiple edges. Hence, we can drop the requirement that the graphs are simple.

A slight variation of Theorem 4 yields a theorem for P_4. Let $C(G)$ denote the number of components of G that consist of C_4.

Theorem 5 *(Hartvigsen [3]) A bipartite graph $G = (V, E)$ has a 4–restricted 2–factor if and only if for every $Z \subseteq V$,*

$$|Z| \geq |A(G-Z)| + |B(G-Z)| + |C(G-Z)|.$$

Necessity is straightforward. The sufficiency is implied by a more general theorem by Király [4] and by my generalization to P_{2k}, the proof of which is given in Section 3.

A few definitions are necessary for the more general statement of Theorem 5. Let $f : V \rightarrow Z^+$ be a function from the nodes of G to the nonnegative integers. For each $v \in V$, we call $f(v)$ the *weight* of v. Because this is a generalization of 2–factors, we consider only the functions f which assign each node a weight of 1 or 2. So let $f : V \rightarrow \{1,2\}$. For a graph $G = (V, E)$, let $I(G)$ denote the set of isolated nodes in G, $J(G)$ the set of two–node components in G with total weight 4, and $K(G)$ the set of C_4 components in G with total weight 8. Finally, let $q(Z) = f(I(G-Z)) + 2|J(G-Z)| + 2|K(G-Z)|$ for $Z \subseteq V$.

Theorem 6 *(Király [4]) Let G be a bipartite graph with bipartition (X,Y) and $f: V \to \{1,2\}$. Then G has a 4-restricted f-factor if and only if*

$$f(Z) \geq q(Z) \quad \forall Z \subseteq V.$$

Thereom 6 is implied by our generalization given in Section 3.

It seems as if Theorem 5 and Theorem 6 should extend to a theorem for P_{2k} in bipartite graphs because Theorem 4 extends to Theorem 5 and Theorem 6. However, it will be shown in the next section that the logical conjecture for P_{2k} in bipartite graphs is only valid for a particular class of bipartite graphs.

In his paper, Király also states an Ore–type deficiency theorem which is easily obtained from Theorem 6. His result is important because it provides a means of evaluating how far a graph is from having a 4–restricted f–factor where $f(v) = 1$ or 2 for every $v \in V$.

Corollary 7 *(Király [4]) Let $G = (V, E)$ be a bipartite graph with bipartition (X,Y) and $f: V \to \{1,2\}$. Let $def_U(G)$ denote the minimum of $f(U) - f'(U)$ where the minimum is taken over all functions f' such that $f': V \to \{1,2\}$, $0 \leq f'(v) \leq f(v)$ for all $v \in V$ and G has a 4-restricted f'-factor. Then*

$$def_X(G) = \frac{\max_{Z \subseteq V}(q(Z) - f(Z)) - f(Y) + f(X)}{2}, \quad \textit{and}$$

$$def_V(G) = \max_{Z \subseteq V}(q(Z) - f(Z)).$$

In the next section, we prove that these Tutte–type results for P_4 in bipartite graphs will generalize to theorems for P_{2k} in a special class of bipartite graphs.

3 P_{2k} in Bipartite Graphs

In order to determine the special class of bipartite graphs, we need to state a theorem concerning Hamilton–connectedness. A bipartite graph G with bipartition (X, Y) is called *Hamilton–connected* if for every $u \in X$ and $v \in Y$ there exists a Hamilton path with u and v as endpoints. G is called *balanced* if $|X| = |Y|$. Let $\delta(G)$ denote the minimum degree of G.

Theorem 8 *(Bagga and Varma [1]) Let G be a balanced bipartite graph of order $2n$ such that $\delta(G) \geq \frac{n+2}{2}$. Then G is Hamilton–connected.*

Let k be given, and recall that $I(G)$ is the set of isolated nodes of G. Let $f : V \rightarrow \{1, 2\}$, and let $\overline{C}_i$ denote a cycle of length i with arbitrarily many chords and total weight $2i$ ($\sum_{v \in \overline{C}_i} f(v) = 2i$). Let $J_1(G)$ be the set of K_2 components of total weight 4 in G and $J_i(G)$ be the set of $\overline{C}_{2i}$ components of G for $1 \leq i \leq k$. For $Z \subseteq V$ define

$$q_i(Z) = f(I(G - Z)) + 2\sum_{j=1}^{i} |J_i(G - Z)|. \tag{1}$$

A subset Z is *violating* when $f(Z) < q_k(Z)$ and *tight* when $f(Z) = q_k(Z)$. Let (X, Y) be the bipartition of G. Then Z is called *nontrivial* if it intersects both X and Y. If k is given, then let B_{2k} denote the set of all bipartite graphs G such if C_{2i} is a subgraph of G, $2 \leq i \leq k$, then $H = G[V(C_{2i})]$ (the graph induced by the nodes of C_{2i}) satisfies $\delta(H) \geq \frac{i+2}{2}$. Hence, H is Hamilton–connected. Now we can state the following generalization of Theorem 6.

Theorem 9 *Let $2k$ be given, and let $G = (V, E)$ be a bipartite graph in $\mathcal{B}_{2k}$ with bipartition (X, Y). Let $f : V \rightarrow \{1, 2\}$. Then G has a $2k$-restricted f-factor if and only if*

$$f(Z) \geq q_k(Z) \quad \forall Z \subseteq V. \tag{2}$$

Proof: The necessity is trivial for any graph. In any f–factor, at least $q_k(Z)$ edges must go from $G - Z$ to Z, and at most $f(Z)$ edges are incident with some node in Z for any $Z \subseteq V$.

We will prove the sufficiency by induction on $f(V)$. First note that it must be the case that $f(X) = f(Y)$. Otherwise, without loss of generality, assume that $f(X) < f(Y)$, and let $Z = X$. Then we have that $f(Z) = f(X)$, and $q_k(Z) = f(Y)$ which contradicts the assumption that (2) holds for every $Z \subseteq V$. Suppose that $f(V) = |V|$ (clearly $f(V) \geq |V|$). Then an f–factor is a perfect matching, and for every $Z \subseteq V$, $J_i(G - Z) = \emptyset$ for $1 \leq i \leq k$. Let $Z \subseteq X$, and let $N(Z)$ denote the set of nodes in $V \backslash Z$ which are adjacent to some node in Z. Then $f(N(Z)) = |N(Z)|$, and $q_k(N(Z)) = |I(G - N(Z))| \geq |Z|$. The Marriage Theorem states that a bipartite graph has a perfect matching if and only if $|X| = |Y|$, and $|Z| \leq |N(Z)|$ for every

$Z \subseteq X$. Hence, G has an f–factor (or perfect matching).

By the induction hypothesis, we now know that if $f' : V \to \{1, 2\}$ and $G' \in \mathcal{B}_{2t}$ is a graph such that $f'(V(G')) < f(V(G))$ and f', G' satisfy (2), then G' has a $2k$–restricted f'–factor.

Claim 1 *If there are no edges between nodes of weight 2, then G has a $2k$–restricted f-factor.*

Proof: For every $u \in V$ such that $f(u) = 1$, let $f'(u) = 1$. For every $v \in V$ such that $f(v) = 2$, split v into v_1 and v_2 and let $f'(v_1) = f'(v_2) = 1$. Let v_i be incident with every node that v is incident with in G. Let $G' = (V', E')$ be the resulting graph ($uv \in E$ and $f(v) = 2 \to uv_1, uv_2 \in E'$). Let $Z' \subseteq V'$. Now define $Z \subseteq V$ by:

(i) if $v \in Z'$ is not an expanded node then let $v \in Z$

(ii) if $v_1, v_2 \in Z'$ then let the contracted node $v \in Z$

(iii) if $v_1, v_2 \notin Z'$ then the contracted node $v \notin Z$

(iv) if exactly one of v_1 and v_2 is in Z', then the contracted node $v \notin Z$.

It follows that $f'(Z') \geq f(Z) \geq q_k(Z) \geq q'_k(Z')$. Hence, G' and f' satisfy (2). G' has no nodes of weight 2 with respect to f'; therefore, by the argument given above, G' has a $2k$–restricted f'–factor (perfect matching). In G' contract each v_1, v_2 back to v to get a $2k$–restricted f-factor in G. ■

We may now assume that G has at least one edge between nodes of weight 2.

Claim 2 *Every $\overline{C}_{2i}$ subgraph of G has at least one node in each partite set with neighbors outside of $\overline{C}_{2i}$ for $2 \leq i \leq k$.*

Proof: Without loss of generality, suppose that there exists a $\overline{C}_{2i}$ subgraph of G such that $N(\overline{C}_{2i}) \cap X = \emptyset$. Let $Z = Y \backslash V(\overline{C}_{2i})$. Then $G - Z$ has $|X| - i$ isolated nodes and one $\overline{C}_{2i}$ component. Hence, $f(Z) = f(Y) - 2i$ and $q_k(Z) = f(X) - 2i + 2$. So we have $f(Z) < q_k(Z)$ which is a contradiction. ■

Because we're using induction on $f(V)$, we need some way of reducing G, f to G', f' such that $f'(V(G')) < f(V(G))$. Hence, the following procedures will be useful.

Edge-reduction(x,y): If $xy \in E$ and $f(x) = f(y) = 2$, then delete the edge xy and let $f'(x) = f'(y) = 1$. Let $f'(u) = f(u) \quad \forall u \in V \backslash \{x, y\}$.

$\overline{C}_{2i}$-reduction($x_1, y_1, ..., x_i, y_i$): If $x_i \in X$, $y_i \in Y$ are the nodes of a $\overline{C}_{2i}$, then

(i) contract $x_1, ..., x_i$ to a new node x and contract $y_1, ..., y_i$ to a new node y

(ii) delete the parallel edges and the edge xy

(iii) let $f'(x) = f'(y) = 1$ and $f'(u) = f(u) \quad \forall u \in V \backslash \{x_1, y_1, ..., x_i, y_i\}$.

The goal is to find a $2k$–restricted f'–factor in a reduced graph, and extend it to a $2k$–restricted f–factor in G. This extension is the reason for the Hamilton–connected condition on $\overline{C}_{2i}$.

Take any edge $uv \in E$ such that $f(u) = f(v) = 2$. Let $H_1^k, ..., H_{t_k}^k$ denote the $\overline{C}_{2k}$ subgraphs of G containing the nodes u, v. If $t_k > 0$, call $\overline{C}_{2k}$–reduction on the nodes of H_1^k. If the resulting graph G_1^k has a $2k$–restricted f'–factor, F', then it can be extended to a $2k$–restricted f–factor of G because H_1^k is Hamilton–connected. Adding a path of length $2k-1$ cannot create a cycle of length $2k$ since the contracted nodes x and y are not adjacent in G_1^k. Figure 1 demonstrates this process. The solid lines are edges in the f–factor, and the dotted lines are other edges in G.

Let G_i^k be the graph obtained by calling $\overline{C}_{2k}$–reduction on the nodes of H_i^k starting from the graph G each time. By the previous argument and the induction hypothesis, if any one of the reduced graphs $G_1^k, ..., G_{t_k}^k$ satisfies (2), then we can find a $2k$–restricted f–factor in G. Therefore, we may assume that each G_i^k has a violating set. Let Z_i^k be a maximal set violating (2) in G_i^k for $1 \le i \le t_k$, and let x and y be the contracted nodes resulting from the reduction of H_i^k.

Claim 3 *If Z is a maximal violating set, then every component of $G - Z$ must be contained in $I(G - Z) \cup (\bigcup_{i=1}^k J_i(G - Z))$.*

Proof: Suppose that C is a component of $G - Z$ that is not contained in $I(G - Z) \cup (\bigcup_{i=1}^k J_i(G-Z))$. Assume without loss of generality that $f(V(C) \cap X) \le f(V(C) \cap Y)$.

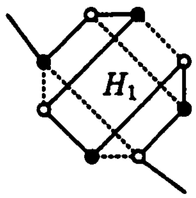


Figure 1 : Extending the f'–factor to an f–factor

Then,

$$f(Z \cup (V(C) \cap X)) = f(Z) + f(V(C) \cap X) < \\ q_k(Z) + f(V(C) \cap Y) = q_k(Z \cup (V(C) \cap X) \tag{3}$$

This contradicts the assumption that Z is a maximal violating set. ■

Claim 4 *Z_i^k must contain the contracted nodes x and y.*

Proof: Without loss of generality, suppose that $y \notin Z_i^k$. By the maximality of Z_i^k, every component of $G_i^k - Z_i^k$ must be an isolated node, a $\overline{C}_4$, $\overline{C}_6$, . . ., or a $\overline{C}_{2k}$. Now we'll construct Z in G. Take Z_i^k and if it contains x, replace x with the nodes in $V(G)$ which were contracted to form x. Likewise, replace y with the nodes in $V(G)$ which were contracted to form y. We know that y is an isolated node in $G_i^k - Z_i^k$ because $f'(y) = 1$. This implies that the nodes in $V(G)$ which were contracted to form y are isolated in $G - Z$. Therefore, $f(Z) = f'(Z_i^k) + 2k - 1 < q_k'(Z_i^k) + 2k - 1 = q_k(Z)$. This implies that Z is a violating set in G which is a contradiction. ■

Now we may assume that there exist maximal nontrivial violating sets $Z_1^k, ..., Z_{t_k}^k$ of $G_1^k, ..., G_{t_k}^k$ respectively, such that $x, y \in Z_i^k$ for $1 \leq i \leq t_k$.

Let $H_1^{k-1}, ..., H_{t_{k-1}}^{k-1}$ be the $\overline{C}_{2k-2}$ subgraphs of G containing the nodes u, v. Call $\overline{C}_{2k-2}$–reduction on the nodes of H_i^{k-1} each time starting from G. Let $G_1^{k-1}, ..., G_{t_{k-1}}^{k-1}$ be the resulting graphs. Suppose that some G_i^{k-1} has a $2k$–restricted f'–factor, F'. Then clearly F' can be extended to a $(2k-2)$–restricted f–factor F of G by adding to F' a path of length $2k-3$. Note that this is due to the fact that H_i^{k-1} is Hamilton connected.

Suppose that F contains a cycle of length $2k$. Then the cycle must be contained in H_j^k for some j. Note that F minus the edges of H_j^k is a $2k$–restricted f'–factor in $G_j^k - x - y$ (x and y are the contracted nodes in G_j^k). However we also have that,

$$f'(Z_j^k - x - y) = f'(Z_j^k) - 2 < q_k'(Z_j^k) - 2 = q_k'(Z_j^k - x - y) - 2. \tag{4}$$

Hence, $Z_j^k - x - y$ is a violating set in $G_j^k - x - y$. A graph with a violating set cannot have a $2k$–restricted f–factor. Therefore, it must be the case that F is a $2k$–restricted f–factor in G. By the induction hypothesis and the previous argument, we may now assume that G_i^{k-1} has a maximal violating set Z_i^{k-1} for $1 \le i \le t_{k-1}$. By the same proof that is given for Claim 4, $x, y \in Z_i^{k-1}$ for $1 \le i \le t_{k-1}$. Hence, Z_i^{k-1} is a maximal nontrivial violating set in G_i^{k-1}.

For $j = k-2, k-3, ..., 2$, let $H_1^j, ..., H_{t_j}^i$ be the $\overline{C}_{2j}$ subgraphs of G containing the edge uv. Call $\overline{C}_{2j}$–reduction on the nodes of H_i^j each time starting from G. Let $G_1^j, ..., G_{t_j}^j$ be the resulting graphs. By the same arguments given above, we may assume that there exist maximal nontrivial violating sets Z_i^j in G_i^j such that $x, y \in Z_i^j$ for $1 \le i \le t_j$ and $j = 2, 3, ..., k$.

Now call edge–reduction on uv, again starting with G. If the resulting graph G' satisfies (2), then by induction G' has a $2k$–restricted f'–factor, F'. Clearly, F' extends to an f–factor F of G. Suppose that F contains a cycle of length $2j \le 2k$.

(Note that F can contain at most one cycle of length $2k$ or less because adding the edge uv can create at most one cycle.) This cycle must be contained in H_i^j for some i. Then F minus the edges of H_i^j is a $2k$–restricted f'–factor in $G_i^j - x - y$. However, since x and y are in the violating set Z_i^j,

$$f'(Z_i^j - x - y) = f(Z_i^j) - 2 < q_k(Z_i^j) - 2 = q(Z_i^j - x - y) - 2. \tag{5}$$

This is a contradiction so we may now assume that G' has a maximal violating set Z' because otherwise we can find a $2k$–restricted f–factor in G.

Let (X', Y') be the bipartition of G'. Note that $f'(X') = f(X) - 1 = f(Y) - 1 = f'(Y)$. This implies that

$$\begin{aligned} &f'(Z' \cap X') + f'((V(G') \backslash Z') \cap X') \equiv \\ &f'(Z' \cap Y') + f'((V(G') \backslash Z') \cap Y') \pmod 2. \end{aligned} \tag{6}$$

Use this parity condition to see that

$$\begin{aligned} &f'(Z' \cap X') + |\{w \in (V(G') \backslash Z') \cap X' : f'(w) = 1\}| \equiv \\ &f'(Z' \cap Y') + |\{w \in (V(G') \backslash Z') \cap Y' : f'(w) = 1\}| \pmod 2. \end{aligned} \tag{7}$$

This implies that

$$f'(Z'\cap X')+f'(Z'\cap Y')\equiv |\{w\in (V(G')\backslash Z')\cap X' : f'(w)=1\}|+ \\ |\{w\in (V(G')\backslash Z')\cap Y' : f'(w)=1\}| \pmod 2 \tag{8}$$

Hence, $f'(Z')\equiv q_k(Z') \pmod 2$. Therefore, since $f'(Z')<q_k(Z')$, we also have that

$$f'(Z')\le q_k'(Z')-2. \tag{9}$$

Let $Z=Z'$ and replace the edge uv to get G. We must show that Z is a tight nontrivial set in G. Suppose that $u,v\notin Z$. Then it must be the case that u and v are isolated nodes in $G'-Z'$ because $f'(u)=f'(v)=1$ and every component of $G'-Z'$ is either an isolated node (with weight 1 or 2) or contained in $\bigcup_{i=1}^{k} J_i(G-Z)$ (each of whose nodes has weight 2). This implies that uv is a weight 4 K_2 in $G-Z$ and

$$f(Z)=f'(Z')<q_k'(Z')=q_k(Z). \tag{10}$$

This contradicts the original assumption that G satisfies (2). Without loss of generality, suppose that $u\in Z$ and $v\notin Z$. Then v is an isolated node in both $G'-Z'$ and $G-Z$ and

$$f(Z)=f'(Z')+1<q_k'(Z')+1=q_k(Z). \tag{11}$$

Again, this contradicts the original asumption that G satisfies (2). So now we can conclude that $u, v \in Z$, and by (9), $f(Z) = f'(Z') + 2 \leq q_k(Z') = q_k(Z)$. Since G satisfies (2), it must be the case that $f(Z) = q_k(Z)$. Hence, Z is a maximal nontrivial tight set.

We have now shown that G either has a $2k$–restricted f–factor or a maximal nontrivial tight set Z.

Recall that each component of $G - Z$ is an isolated node, a weight 4 K_2, or a $\overline{C}_{2i}$ for some $2 \leq i \leq k$. Call edge–reduction on each K_2 component and $\overline{C}_{2i}$–reduction on each $\overline{C}_{2i}$ component for $2 \leq i \leq k$. In the resulting graph, delete all edges between nodes in Z to obtain the bipartite graph G' with bipartition (X', Y'). G' can clearly be separated into two parts: G_1 which is spanned by $(Z \cap X') \cup (Y' \backslash Z)$ and G_2 which is spanned by $(Z \cap Y') \cup (X' \backslash Z)$ (see Figure 2). No edge joins a node in G_1 to a node in G_2.

We know that Z is non–trivial, so $Z \cap X' \neq \emptyset$ and $Z \cap Y' \neq \emptyset$ which implies that $f(V(G_i)) < f(V(G))$ for $i = 1, 2$. Note that the reductions did not change $q_k(Z)$, so Z is also tight in G'. Since $f(X) = f(Y)$, after the reductions we have $f'(X') = f'(Y')$. Let V' denote $V(G')$. Every component of $G' - Z$ is an isolated node, and Z is tight so $f'(Z) = q'_k(Z) = f'(V' \backslash Z)$. Therefore, we have:

$$f'(Z \cap X') + f'(Z \cap Y') = f'(X' \backslash Z) + f'(Y' \backslash Z) \qquad (12)$$

$$f'(Z \cap X') + f'(X' \backslash Z) = f(Z \cap Y') + f'(Y' \backslash Z). \qquad (13)$$

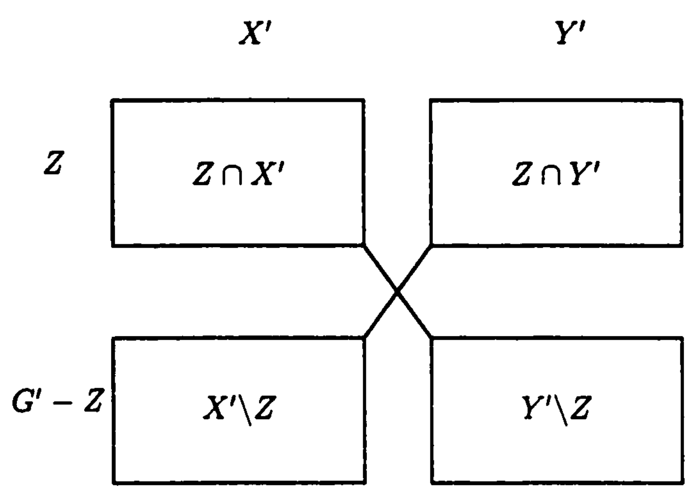


Figure 2 : G' with bipartition (X', Y')

These equalities imply that $f'(Z \cap X') = f'(Y' \backslash Z)$ and $f'(Z \cap Y') = f'(X' \backslash Z)$. These facts will be useful in proving the following claim.

Claim 5 *G_1 and G_2 satisfy (2).*

Proof: Suppose that Z_1 is a violating set in G_1. Let $Z' = Z_1 \cup (Z \cap Y')$. Note that $Z_1 \cap (Z \cap Y') = \emptyset$ and $q_k(Z \cap Y') = f'(X' \backslash Z)$ in G_2. Therefore, in G',

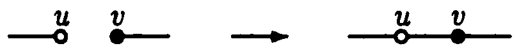
$$\begin{aligned} f'(Z') = & f'(Z_1) + f'(Z \cap Y') < q'_k(Z_1) + f'(Z \cap Y') = \\ & q'_k(Z_1) + f'(X' \backslash Z) = q'_k(Z_1) + q'_k(Z \cap Y') = q'_k(Z'). \end{aligned} \tag{14}$$

Hence, Z' is violating in G'. Let $\overline{Z} \subseteq V(G)$ be the set of nodes obtained from Z' by replacing every contracted node in Z' by its ancestors. The contracted nodes in Z'

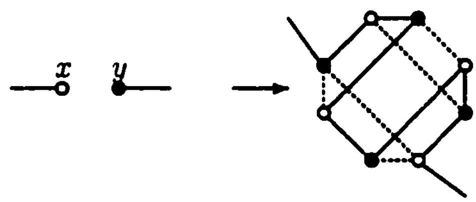
are the contracted nodes from Y', so in each reversed reduction step, $f(\overline{Z})$ and $q_k(\overline{Z})$ are increased by the same amount. This implies that $\overline{Z}$ is a violating set in G which is a contradiction. By the same argument, G_2 also satisfies (2). ■

Therefore, G_1 and G_2 satisfy (2). Let F_i be a 2k-restricted f'-factor in G_i for $i = 1, 2$. Clearly $F' = F_1 \cup F_2$ is a $2k$-restricted f'-factor in G'. Define F by F' along with:

(i) all reduced edges



(ii) a path of length $2i - 1$ for each reduced $\overline{C}_{2i}$, $1 \leq i \leq k$



Clearly F is an f-factor of G. It remains to be shown that F has no cycles of length $2k$ or less.

Suppose that F contains some cycle C_{2i} where $2 \leq i \leq k$. Then let $H = G[V(C_{2i})]$ (the subgraph of G induced by $V(C_{2i})$). Since F' is a $2k$-restricted f'-factor, C_{2i} must have been created by the addition of one or more paths. It's easy to see that there must be an even number of such additions (see Figure 3).

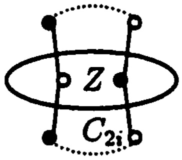


Figure 3 : Even number of added paths in C_{2i}

Let q be the number of added paths, so q is even and nonzero. Let l be the length of the shortest added path P, so $l \geq 1$ and of odd parity. Let $n_1 = |I(G - Z) \cap X \cap V(C_{2i})|$, the number of isolated nodes of $V(C_{2i})$ in $V(G - Z) \cap X$, and $n_2 = |I(G-Z) \cap Y \cap V(C_{2i})|$, the number of isolated nodes of $V(C_{2i})$ in $V(G-Z) \cap Y$. Along the cycle, each node counted in n_1 must have two neighbors in $V(C_{2i}) \cap Z \cap Y$, and, similarly, each node counted in n_2 must have two neighbors in $V(C_{2i}) \cap Z \cap X$. For each added path, we must have an endpoint with a neighbor in $V(C_{2i}) \cap Z \cap X$ and an endpoint with a neighbor in $V(C_{2i}) \cap Z \cap Y$. In $V(C_{2i})$, each node in Z must have exactly 2 neighbors in $V(C_{2i}) \backslash Z$ (see Figure 4). This implies the following:

$$|Z \cap X \cap V(C_{2i})| = \frac{q}{2} + n_2 \tag{15}$$

$$|Z \cap Y \cap V(C_{2i})| = \frac{q}{2} + n_1. \tag{16}$$

Let $v_1 \in V(P) \cap X$ and $v_2 \in V(P) \cap Y$ (see Figure 4). Note that by the definition of $\mathcal{B}_{2k}$, $deg_H(v_j) \geq \frac{i+2}{2}$ for $j = 1, 2$. Also, in H, v_1 can only have neighbors in the set $V(C_{2i}) \cap P \cap Y$, and v_2 can only have neighbors in the set $V(C_{2i}) \cap P \cap X$. This implies that $n_1 + \frac{q}{2} + \frac{l+1}{2} \geq deg_H(v_1) \geq \frac{i+2}{2}$, and $n_2 + \frac{q}{2} + \frac{l+1}{2} \geq deg_H(v_2) \geq \frac{i+2}{2}$. We can simplify these inequalities to get:

$$2n_1 + q + l \geq i + 1 \tag{17}$$

$$2n_2 + q + l \geq i + 1 \tag{18}$$

Now find a lower bound on $2i$, the number of nodes in C_{2i}, by adding the nodes in $Z \cap V(C_{2i})$, plus the nodes in $I(G - Z) \cap V(C_{2i})$, plus a lower bound on the number of nodes which are on added paths. Note that $q \geq 2$.

$$\begin{aligned} 2i &\geq n_2 + \frac{q}{2} + n_1 + \frac{q}{2} + n_1 + n_2 + q(l+1) \\ &\geq 2n_1 + 2n_2 + 2q + 2l \\ &\geq 2(i+1) = 2i + 2 \end{aligned} \tag{19}$$

This is a contradiction. Hence, F has no cycle of length $2k$ or less and is therefore a $2k$-restricted f-factor of G. ■■

We have now proven a necessary and sufficient result for a particular class of bipartite graphs. It remains to be shown why Theorem 9 does not hold for every bipartite graph. Let $k \geq 3$ and consider the following graph. Let $f' : V(G') \rightarrow \{1, 2\}$

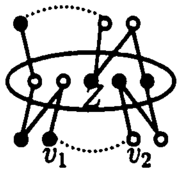


Figure 4 : C_{2i} with v_1 and v_2

where G' is any bipartite graph with bipartition (X', Y'), $X', Y' \neq \emptyset$, which has a $2k$–restricted f'–factor; hence, G' satisfies (2). Consider the graph C_6 and let x and y be two nonadjacent nodes of C_6 that are in different partite sets. To obtain G, add an edge between x and every node in Y', and add an edge between y and every node in X' (see Figure 5). Define f by $f(v) = f'(v)$ if $v \in V(G')$ and $f(v) = 2$ if $v \in V(C_6)$.

Theorem 10 *The graph G described above satisfies (2) but has no $2k$–restricted f–factor.*

Proof: First we must show that G has no $2k$–restricted f–factor. Each node v in $V(C_6) - x - y$ has degree 2 and $f(v) = 2$, so any f–factor must use every edge of C_6. Hence, G has no $2k$–restricted f–factor.

Let Z be a maximal violating set in G ($f(Z) < q_k(Z)$). If $Z \cap V(C_6) = \emptyset$, then Z is a violating set in G'. This contradicts the assumption that G' has a $2k$–restricted f'–factor. Therefore, $Z \cap V(C_6) \neq \emptyset$. Also, if $V(C_6) \subseteq Z$, then $Z \backslash V(C_6)$ is a violating set in G'. Hence $\emptyset \neq Z \cap V(C_6) \neq V(C_6)$. Since G' has no violating set, each node

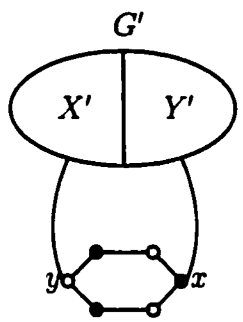


Figure 5 : $G = G' \cup C_6$

of $V(C_6)\backslash Z$ must appear in a component of $G - Z$ that is either an isolated node, a single edge, or the entire C_6. The only way for C_6 to be a component of $G - Z$ is to let $Z = V(G')$. But then Z is not violating because $|Z| \geq 2 \implies f(Z) \geq 2$ and $q_k(Z) = 2$. The only other option is for $V(C_6) \cap Z = \{x, y\}$. Let $Z' = Z - x - y$. Then refer to Figure 5 to see that:

$$f'(Z') = f(Z) - 4 < q_k(Z) - 4 = q'_k(Z') \tag{20}$$

This contradicts the assumption that G' has a $2k$-restricted f'-factor. ■

It may be possible to improve Theorem 9 by finding a better minimum degree condition for the $\overline{C}_{2i}$ subgraphs to be Hamilton-connected. It can be improved for the case where $k = 3$. Figure 6 demonstrates that $\overline{C}_6$ with two chords is Hamilton-connected. The solid lines are edges of the Hamilton path, and the dotted lines are other edges in $\overline{C}_6$. Hence, we have the following theorem:

Theorem 11 *Let $G = (V, E)$ be a bipartite graph with bipartition (X, Y) such that every $\overline{C}_6$ subgraph of G contains at least two chords. Let $f : V \to \{1, 2\}$. Then G has a 6-restricted f-factor if and only if*

$$f(Z) \geq q_3(Z) \quad \forall Z \subseteq V.$$

Clearly given Figure 6, Theorem 11 follows directly from Theorem 9. Hence, we know that Theorem 9 does not hold for every bipartite graph, but it does hold for some bipartite graphs that are not included in $\mathcal{B}_{2k}$.

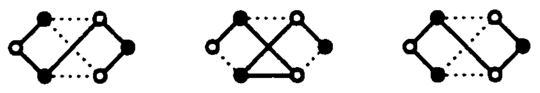


Figure 6 : Hamilton paths in C_6 with two chords

In the remaining sections we consider additional results for P_4 in bipartite graphs. These may generalize to results for P_6 or P_{2k} in bipartite graphs.

4 Augmenting Structures

Let $G = (V, E)$ be a simple graph. A *simple 2-matching in G* is a subgraph H of G each of whose components is a simple path or a cycle. Hence the degree of each node in H is 1 or 2, and H is not necessarily spanning. The *cardinality* of H is the number of edges in H. A *4-restricted simple 2-matching* is a simple 2-matching with no cycles of length 4 or less. We will consider the problem of determining a maximum cardinality 4-restricted simple 2-matching in a bipartite graph.

Hartvigsen discovered an algorithm for finding maximum cardinality 4-restricted simple 2-matchings in bipartite graphs that is similar to Edmonds' algorithm for finding a maximum matching (see [3]). The matching algorithm grows a matching by iteratively finding augmenting paths (a path with exposed endpoints that alternates between edges in the matching and edges not in the matching). A similar idea is used for 4-restricted simple 2-matchings.

Let $G = (V, E)$ be a graph and M a simple 2-matching. A node in V is called *saturated* if it is incident with two edges of M. If it is incident with one or zero edges of M, then it is called *deficient*. An *augmenting path* with respect to M is defined to be a simple path such that:

(i) the edges are alternately in and out of M

(ii) the endpoints are deficient

(iii) the endedges are not in M.

Clearly a maximum simple 2–matching cannot have an augmenting path. It turns out that the converse of this statement is true as well. Hence we have the following theorem.

Theorem 12 *(Hartvigsen [3]) A simple 2–matching M in a bipartite graph G has maximum cardinality if and only if G has no augmenting path with respect to M.*

In order to state the corresponding theorem for 4–restricted simple 2–matchings, we need something similar to an augmenting path. Let M be a 4–restricted simple 2–matching. An *augmenting structure* with respect to M is a simple path P with a possibly empty collection of cycles $C_1, ..., C_p$ such that:

(i) $P, C_1, ..., C_p$ are pairwise edge–disjoint

(ii) P is an augmenting path with respect to M

(iii) in each C_i, the edges are alternately in and out of M

(iv) interchanging the edges of $P, C_1, ..., C_p$ in and out of M results in a 4–restricted simple 2–matching.

In Figure 7 we have an example of an augmenting structure. The solid lines are edges in M and the dotted lines are other edges in the graph. The path from a to f together with C_1 and C_2 is an augmenting structure. Interchanging the edges along the path and the edges along the two cycles will create a 4–restricted simple 2–matching M' such that $|M'| = |M| + 1$. Clearly, a maximum 4–restricted simple 2–matching cannot have an augmenting structure. This was proven to be a sufficient condition by Hartvigsen.

Theorem 13 *(Hartvigsen [3]) A 4–restricted simple 2–matching M in a bipartite graph G has maximum cardinality if and only if G contains no augmenting structure with respect to M.*

An algorithm given by Hartvigsen (see [3]) uses this theorem to find 4–restricted sim-

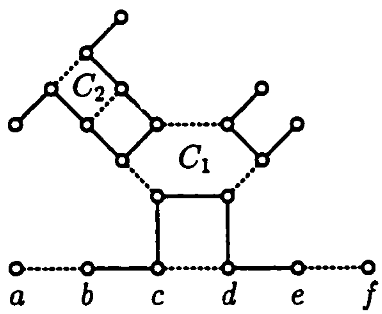


Figure 7 : An augmenting structure

ple 2–matchings of maximum cardinality in bipartite graphs. It is similar to Edmonds' matching algorithm. The algorithm iteratively searches for augmenting structures using a complex tree structure. The author, Hartvigsen, has not yet published a paper on the procedure, but it is similar to his algorithm for 2–matchings in bipartite graphs which is given below.

Recall that a deficient node is a node that is incident with one or zero edges of M. The algorithm will maintain a structure S that is similar to an alternating tree. The nodes of S are either *odd* or *even*. Odd nodes must be incident with two edges of M; however, these edges are not necessarily in S. Even nodes may be saturated or deficient.

INPUT: bipartite graph $G = (V, E)$, simple 2–matching M (i.e. $M = \emptyset$)

OUTPUT: maximum cardinality simple 2–matching

STEP 1: If M is a 2–factor, then we're done. Otherwise, let S be the set of deficient nodes, where each node is even and is a root.

STEP 2: Select an edge $e \notin M$ joining an even node v to a node w that is not odd. If no such edge exists, then M has maximum cardinality.

CASE 1: $w \notin S$ In this case we can grow the structure. Clearly w is incident with two edges, wu_1 and wu_2, in M (otherwise $w \in S$). Neither u_1 nor u_2 is odd, because then w would be even. Hence each u_i is either even or not in S. Make w odd and add it along with the edge vw to the tree that contains v. If $u_i \notin S$, then make it even and add both u_i and the edge wu_i to the tree that contains v. If u_i is even, then don't change its status. Go to STEP 1.

CASE 2: w is even If v and w are in the same tree, then there is an odd cycle. However, G is bipartite, so v and w must be in different trees. Let r_v and r_w be the roots of the trees containing v and w respectively. The path from $r_v ... vw ... r_w$ is an augmenting path. Interchange the edges in and out of the matching along this path to obtain a larger simple 2-matching M'. Go to STEP 1.

The algorithm for the bipartite case of 4-restricted 2-matchings is much more difficult to state carefully; however, it is quite similar to the algorithm given above for 2-matchings in bipartite graphs.

5 Polyhedral Structure

In his paper [3], Hartvigsen presented polyhedral results for simple 2-matchings and for 4-restricted simple 2-matchings. Let $G = (V, E)$ be a graph, and for $v \in V$, let $\delta(v)$ denote the edges in E that are incident with v. For a simple 2-matching M, let $x \in \mathbf{R}^{\mathbf{E}}$ be the incidence vector for M (i.e. $x_e = 1$ if $e \in M$ and $x_e = 0$ if $e \notin M$). For $S \subseteq E$, let $x(S)$ denote $\sum_{e \in S} x_e$. Consider the following linear program (P):

$$\max x(E) \quad s.t.$$

$$x(\delta(v)) \leq 2 \quad \forall v \in V$$

$$0 \leq x_e \leq 1 \quad \forall e \in E.$$

Note that any integral solution to (P) must be the incidence vector of a simple 2–matching. Therefore, an optimal integral solution is maximum cardinality simple 2–matching. When G is bipartite, the constraint matrix for (P) is totally unimodular. This implies the following theorem:

Theorem 14 *(Hartvigsen [3]) For any bipartite graph $G = (V, E)$, (P) has an integer optimal solution.*

Now consider the following variant of (P) which we denote (P'). Let $\mathcal{A} = \{S \subseteq V : S$ is the edge set of a square $\}$.

$$\begin{aligned} \max x(E) &\quad s.t. \\ x(\delta(v)) \leq 2 &\quad \forall v \in V \\ x(S) \leq 3 &\quad \forall S \in \mathcal{A} \\ 0 \leq x_e \leq 1 &\quad \forall e \in E \end{aligned}$$

Note that any integral solution to (P') is a 4–restricted simple 2–matching. Hence, an optimal integral solution to (P') is a maximum cardinality 4–restricted simple 2–matching. Unfortunately, the constraint matrix for (P') is not totally unimodular in general. However, Hartvigsen did obtain the following result:

Theorem 15 *(Hartvigsen [3]) For a bipartite graph $G = (V, E)$, (P') has an integer optimal solution.*

Hence, there exist polynomial time algorithms to find both simple 2–matchings and 4–restricted simple 2–matchings.

There are no results for the polyhedral structure of weighted P_k in bipartite graphs. however, some work has been done on the polyhedral structure of weighted P_k in general graphs. By viewing the TSP as a special case of weighted P_k, some additional inequalities have been found for this well–known problem. Polyhedral results in the weighted case of restricted 2–factors in general graphs have been obtained by Cunningham and Wang (see [2]). In their paper, they introduce a large class of valid inequalities, called *bipartition inequalities*. These inequalities contain both the subtour and comb inequalities as special cases. For $k = 3$ (P_3, or triangle–free 2–factors), they derive a necessary and sufficient condition for these inequalities to be facet inducing. However, it is also shown that the bipartition inequalities are not sufficient to describe the optimal 3–restricted 2–factor polytope. Hence, there is still some work to be done on weighted P_3 in general graphs.

Bibliography

[1] K.S. Bagga and B.N. Varma, *Bipartite graphs and degree conditions*, Graph Theory, Combinatorics, Algorithms, and Applications (Y. Alavi, F.R.K. Chung, R.L. Graham, and D.F. Hsu, eds.), SIAM, Philadelphia, PA, 1991, pp. 564–573.

[2] W.H. Cunningham and Y. Wang, *Restricted 2-factor polytopes*, Math. Programming, To appear.

[3] D. Hartvigsen, *The square-free 2-factor problem in bipartite graphs*, Integer Programming and Combinatorial Optimization, 7th IPCO Conference Graz **LNCS 1610** (1999), 234–241.

[4] Z. Király, *C_4-free 2-factors in Bipartite Graphs*, (1999), Working paper.

[5] L. Lovasz and M.D. Plummer, *Matching theory*, Akadémiai Kiadó, Budapest, 1986.

[6] M. Rosenfeld, *Research problems*, Discrete Math. **45** (1983), 137–139.

[7] W.T. Tutte, *The factors of graphs*, Canad. J. Math. **4** (1952), 314–328.