JOURNAL OF COMBINATORIAL THEORY, Series A 43, 196–204 (1986)

Minimal Non-Two-Colorable Hypergraphs and Minimal Unsatisfiable Formulas

RON AHARONI

Department of Mathematics, Technion-I.I.T., Haifa 32000, Israel

AND

NATHAN LINIAL

Institute for Mathematics and Computer Science, The Hebrew University, Jerusalem, Israel

Communicated by the Managing Editors

Received December 1, 1985

Seymour (*Quart. J. Math. Oxford* **25** (1974), 303–312) proved that a minimal non 2-colorable hypergraph on n vertices has at least n edges. A related fact is that a minimal unsatisfiable CNF formula in n variables has at least $n+1$ clauses (an unpublished result of M. Tarsi.) The link between the two results is shown; both are given infinite versions and proved using transversal theory (Seymour's original proof used linear algebra). For the proof of the first fact we give a strengthening of König's duality theorem, both in the finite and infinite cases. The structure of minimal unsatisfiable CNF formulas in n variables containing precisely $n+1$ clauses is characterised, and this characterization is given a geometric interpretation.

I. PRELIMINARIES

A bipartite graph $\Gamma=(U, K)$ with bipartition $U=X\cup Y$ will be denoted by $\Gamma=(X, Y. K)$. If $F\subseteq K$, $a\in U$ and $A\subseteq U$ we write $F\langle a\rangle=\{u\in U\colon \{a, u\}\in F\}$, $F(a)$ is the single element of $F\langle a\rangle$ if $|F\langle a\rangle|=1$ and $F[A]=\bigcup\{F\langle a\rangle\colon a\in A\}$. A *matching* in Γ is a subset F of Γ such that $|F\langle u\rangle|\leqslant 1$ for every $u\in U$. If F is a matching, $A=F[X]\subseteq W\subseteq Y$ and $B=F[Y]\subseteq Z\subseteq X$, we say that F is a matching from A into Z and that it is a matching from B into W. If $A\subseteq U$ and $F[U]=A$ for some matching F then A is said to be *matchable.*

If $A\subseteq X$ then a 1-*transversal* of A is a subset T of K such that $T[Y]=A$, $|T\langle a\rangle|=2$ for every $a\in A$, and $|T[C]|\geqslant |C|+1$ for every non-empty subset C of A. (T can be viewed as a function from A into $[Y]^2$, the set of

0097-3165/86 $3.00

subsets of Y of size 2, whose image is a forest, i.e., a circuitless graph.) Lovász characterized those bipartite graphs in which one side has a 1-transversal, as follows:

THEOREM L [4]. *The side X in a finite bipartite graph $\Gamma=(X, Y, K)$ has a 1-transversal if and only if for every non-empty subset C of X there holds:* $|K[C]| \geqslant |C|+1$.

It is easily seen (directly, or using Theorem L and Hall's theorem) that if a subset A of X has a 1-transversal then it is matchable.

A subset C of X is called *critical* if it is matchable, but for every matching F from C into Y there holds $F[C]=K[C]$.

In [2] an extension of Theorem L was given for infinite bipartite graphs, and from it there was derived:

THEOREM AK [2, Corollary 1b]. *The side X in a bipartite graph $\Gamma=(X, Y, K)$ has a 1-transversal if and only if*

(a) *X is matchable and*

(b) *X contains no nonempty critical set.*

A *cover* in a graph $G=(V, E)$ is a set of vertices such that every edge is incident with at least one of them.

A hypergraph $H=(V, E)$ is said to be *2-colorable* if there exists a 2-coloring of V such that every edge contains vertices of both colors. It is *minimal non-2-colorable* if it is non-2-colorable but deleting any edge from E results in a 2-colorable hypergraph. With any hypergraph $H=(V, E)$ we associate a bipartite graph $\Gamma_H=(E, V, K)$, where $\{e, v\} \in K$ iff $v \in e$.

A formula F in the variables x_α is said to be in *conjunctive normal form* (CNF) if $F=\bigwedge\{c_i : i \in I\}$, where $c_i=\bigvee\{x_\alpha : a \in A_i\} \vee \bigvee\{\bar{x}_\beta : \beta \in B_i\}$ for each $i \in I$. The c_i's are the *clauses.* $A_i \cap B_i \neq \varnothing$ is possible. F is *satisfiable* if there is an assignment of truth values so that all the clauses c_i have value 1. The variables x_α, $\alpha \in A_i$ are said to *appear positively* in c_i, and x_β, $\beta \in B_i$ *appear negatively* in c_i. The variables of both types are said to *appear* in c_i. We denote the set of variables of F by V_F and its set of clauses $\{c_i : i \in I\}$ by C_F. We associate with F a bipartite graph $\Gamma_F=(C_F, V_F, K_F)$, where $\{c, x\} \in K_F$ if x appears in c.

A CNF formula F is said to be *minimal unsatisfiable* if it is unsatisfiable, but $\bigwedge C'$ is satisfiable for every proper subset C' of C_F. It is said to be *strongly minimal unsatisfiable* if it is minimal unsatisfiable and for any clause $c \in C_F$ and variable x not appearing in c, adding x or (adding $\bar{x}$) to c makes F satisfiable.

II. A Strong Version of König's Theorem

König's theorem states that in any finite bipartite graph the minimal cardinality of a cover equals the maximal cardinality of a matching. This is easily seen to be equivalent to a version which was proved in [1] to hold also for infinite graphs:

THEOREM K. *In any bipartite graph $\Gamma = (X, Y, K)$ there exists a cover $C = A \cup B$, where $A \subseteq X$ and $B \subseteq Y$, such that A is matchable into $Y \setminus B$ and B is matchable into $X \setminus A$.*

So it turns out that if we give up the symmetry between X and Y the theorem can be strengthened to

THEOREM 1. *In any bipartite graph $\Gamma = (X, Y, K)$ there exists a cover $C = A \cup B$, where $A \subseteq X$, $B \subseteq Y$, such that B has a matching into $X \setminus A$ and A has a 1-transversal into $Y \setminus B$.*

Proof. For every subset Z of Y define $D_\Gamma(Z) = D(Z) = \{x \in X: K\langle x \rangle \subseteq Z\}$. Let Ⓑ be the set of subsets Z of Y having a matching into $D(Z)$. Suppose that $\langle Z_i, i < \kappa \rangle$, is an ascending continuous chain of sets in Ⓑ, and let M_i be the matching of Z_i into $D(Z_i)$. Then, for each i, there holds $M_{i+1}[Z_{i+1} \setminus Z_i] \cap M_i[Z_i] = \emptyset$. Hence $I = M_0 \cup \bigcup_{i<\kappa} M_{i+1} \restriction (Z_{i+1} \setminus Z_i)$ is a matching, which matches $\bigcup_{i<\kappa} Z_i$ into $D(\bigcup_{i<\kappa} Z_i)$, and thus $\bigcup_{i<\kappa} Z_i \in$ Ⓑ. By Zorn's lemma it follows that Ⓑ has a maximal element B.

It suffices to show that there exists a 1-transversal from $A = X \setminus D(B)$ into $Y \setminus B$, since $A \cup B$ is, by the definition of $D(B)$, a cover. Suppose that no such 1-transversal exists. Let $\Gamma' = (A, Y \setminus B, K')$ be the subgraph of Γ spanned by $A \cup (Y \setminus B)$. It suffices to show that in Γ' there is a nonempty subset Z of $Y \setminus B$ which is matchable into $D_{\Gamma'}(Z)$. For, then $B \cup Z$ is matchable in Γ into $D_\Gamma(B) \cup D_{\Gamma'}(Z) = D_\Gamma(B \cup Z)$, contradicting the maximality of B in Ⓑ.

For completeness, let us discuss separately the case that Γ' is finite (although this case is covered also by the argument in the general case.) Since A does not have a 1-transversal in Γ', by Theorem L there exists a nonempty subset C of A such that $|K'[C]| \leqslant |C|$. Take such a C with minimal cardinality. Then, clearly, $|K'[C]| = |C|$, and $|K'[S]| > |S|$ for every non-empty subset S of C. By Hall's theorem it follows that C has a matching I into $Y \setminus B$. Since $|B'[C]| = |C|$ there holds $K'[C] = I[C]$, hence $C \subseteq D_{\Gamma'}(I[C])$, and thus taking $Z = I[C]$ proves the required assertion.

Consider now the general case, i.e., when Γ is possibly infinite. By Theorem AK either A is not matchable or it contains a nonempty critical set C. In the first case the existence of the set Z with the required properties is given by Theorem K. In the second case $Z = K[C]$ satisfies the required conditions.

III. Minimal Non-2-Colorable Hypergraphs and Minimal Unsatisfiable CNF Formulas

Seymour [6] proved that if a hypergraph $H = (V, E)$ is minimal non-2-colorable and $V = \bigcup E$ then $|E| \geqslant |V|$. His proof used linear algebra. We present here an infinite version of this theorem, as well as a new proof.

THEOREM 2. *Let $H = (V, E)$ be a hypergraph such that $V = \bigcup E$. If it is minimal non-2-colorable then there exists a matching in Γ_H from V into E.*

Proof. Apply Theorem 1 to Γ_H and let $A \subseteq E$ and $B \subseteq V$ be as in the theorem. It suffices to show that $B = V$. Suppose it is not the case. Since $A \cup B$ is a cover and $\bigcup E = V$, there must hold $A \neq \emptyset$. The set A of edges has a 1-transversal in $V \backslash B$, and as remarked above this 1-transversal can be viewed as a forest. Since a forest is 2-colorable, it follows that the vertices of $V \backslash B$ can be 2-colored so that no edge in A is monochromatic. Since $A \backslash B$ is a cover all edges in $E \backslash A$ are contained in B, and by the minimality of H and the fact that $V \backslash B \subsetneqq V$ it follows that the elements of B can be 2-colored so that no edge in $E \backslash A$ is monochromatic. Thus H is 2-colorable, a contradiction.

The following is an extension of the theorem to the infinite case:

THEOREM 3. *Let F be a (possibly infinite) CNF formula.*

(a) *If there exists a matching in Γ_F from C_F into V_F then F is satisfiable.*

(b) *If F is minimal unsatisfiable then there exists a matching from V_F into C_F.*

Proof. (a) Suppose that there exists a matching I from C_F into V_F. Then one can assign a truth value to each variable $I(c)$ so as to make c true. Since $I(c_1) \neq I(c_2)$ for $c_1 \neq c_2$, this can be done for each clause c independently and then F is satisfied.

(b) Apply Theorem K to Γ_F, and let $A \subseteq C_F$ and $B \subseteq V_F$ as there. The proof will be complete if we show that $B = V_F$. Suppose that $B \neq V_F$. Let $G = \bigwedge (C_F \backslash A)$. Since $A \cup B$ is a cover, $V_G = B \neq V_F$ and thus $G \neq F$, i.e., $A \neq \emptyset$. By the minimality of F there exists an assignment of truth values to the variables in B which satisfies G. Let I be a matching of A into $V_F \backslash B$.

For each $c \in A$ assign a truth value to $I(c)$ which makes c true. This satisfies the entire formula F.

A result closely related to Seymour's is due to Tarsi. He proved [7] that if F is a finite minimal unsatisfiable CNF formula then $|C_F| \geqslant |V_F| + 1$. Clearly, the above theorem implies Tarsi's result.

This result can, in fact, be derived from Seymour's theorem, in the following way. Let $H = (V, E)$ be a hypergraph defined as follows. Let $V = \{x: x \in V_F\} \cup \{\bar{x}: x \in V_F\} \cup \{f\}$, where f is a new symbol. For each clause c in C_F let $e(c)$ be the set containing f and every variable appearing in c, taken with its sign (thus, for example, if $c = x_1 \vee \bar{x}_2$ then $e(c) = \{f, x_1, \bar{x}_2\}$). Define $E = \{e(c): c \in C_F\} \cup \{\{x, \bar{x}\}: x \in V_F\}$. Then F is satisfiable if and only if H is 2-colorable. To see this, assume that H is 2-colorable, and let V be properly colored red and blue. Suppose, for example, that f is colored red. Since $\{x, \bar{x}\} \in E$, precisely one of $x, \bar{x}$ is colored blue for each $x \in V_F$. Assign x a true value if x is colored blue, and false otherwise. Then, since each clause contains a blue vertex, each clause is satisfied. In the other direction, if there is a truth assignment satisfying F, coloring each vertex x blue if it is true and red if false, and coloring $\bar{x}$ in the opposite color, properly colors H. It is also easy to see that if F is minimal unsatisfiable then H is minimal non 2-colorable. Therefore, by Seymour's result, $|E| = |C_F| + |V_F| \geqslant |V| = 2|V_F| + 1$, hence $|C_F| \geqslant |V_F| + 1$. (The above transformation is taken from [4].)

We also give a linear algebraic proof, analogous to Seymour's proof.

Let M be the matrix indexed by $V_F \times C_F$, where $m_{xc} = 1, -1$ or 0 according to whether x_i appears positively, negatively, or not at all in c. Part (b) will clearly follow if we prove that the rows of M are linearly independent. Suppose that they are dependent, and let $\sum_{x \in V_F} \alpha_x M_x$ be a nontrivial zero linear combination of the rows M_x of M. Let $I_1 = \{x: \alpha_x > 0\}$, $I_2 = \{x: \alpha_x < 0\}$ and $I_3 = \{x: \alpha_x = 0\}$. By the minimality of F the formula $G \wedge D_{\Gamma_F}(I_3)$ is satisfiable, so choose truth values for the variables in I_3 so as to satisfy it. Put $x = $ true for every $x \in I$, and $x = $ false for $x \in I_2$. If $c \notin D_{\Gamma_F}(I_3)$ then at least one term in the sum $\sum \alpha m_{xc}$ is positive, since this sum is zero, and not all of its terms are zero. But by the definition of m_{xc} this means that the above assignment satisfies F, a contradiction.

IV. The Structure of Strongly Minimal Unsatisfiable CNF Formulas F with $|V_F| + 1$ Clauses

We have seen that a minimal unsatisfiable CNF formula with n variables has at least $n + 1$ clauses. It is natural to ask what possible structure such a formula may have if it has exactly $n + 1$ clauses. In this section we solve a special case of this problem by giving a complete description of such formulas which are "strongly minimal." We show that, if F is a strongly minimal formula with n variables and $n+1$ clauses, then there is a variable x which appears in each clause of F so that we may write $F = F_1 \wedge F_2$, where x appears positively in each clause of F_1 and negatively in each clause of F_2. We show further that the formula F_i' $(i = 1, 2)$ obtained by deleting x from F_i, is of the same kind (or is empty) and so has a variable common to all its clauses. Continuing we see that the formula F has the structure of a tree on n nodes whose leaves are formulas of the form $y \wedge \bar{y}$. Conversely, every formula that can be obtained in this manner is strongly minimal and has $n+1$ clauses.

Let us introduce the following notation: if $x \in V_F$ we write C_x, C_x^+, C_x^-, and C_x^0 for the sets of clauses which contain x, contain x positively, contain x negatively, and which do not contain x at all. We write D_x^+ for the set of clauses obtained from clauses in C_x^+ by deleting x from them. A similar definition holds for D_x^-. Note that here we allow empty clauses. Let $F_x^+ = \bigwedge(D_x^+ \cup C_x^0)$ and $F_x^- = \bigwedge(D_x^- \cup C_x^0)$. Also write $V_x^+ = V_{F_x^+}$ and $V_x^- = V_{F_x^-}$.

THEOREM 4. *Let F be a strongly minimal unsatisfiable finite CNF formula such that $|C_F| = |V_F| + 1$. Then there exists a variable x such that*

(a) *x appears in all clauses of F,*

(b) *$V_x^+ \cap V_x^- = \varnothing$, and*

(c) *F_x^+ and F_x^- are strongly minimal unsatisfiable and $|C_x^+| = |V_x^+| + 1$, $|C_x^-| = |V_x^-| + 1$.*

Proof. The formula F_z^+ is unsatisfiable for any $z \in V_F$, since otherwise adding $z = \text{false}$ to the assignment of truth values which satisfies it would satisfy F. It is also minimal unsatisfiable. For, suppose that deleting a clause c from it results in an unsatisfiable formula. If $c \in D_z^+$ then deleting $c \vee z$ from F yields an unsatisfiable formula, contradicting the minimality of F. If $c \in C_z^0$ then replacing c by $c \vee \bar{z}$ in F gives an unsatisfiable formula, contradicting the strong minimality of F. Similarly F_z^- is minimal unsatisfiable.

Define a relation $<$ on V_F by: $y < x$ if $C_y \subseteq C_x^+$ or $C_y \subseteq C_x^-$. Clearly $<$ is transitive. It is also anti-reflexive, since $x < x$ means that either x appears only positively in F or it appears only negatively. But then by the minimality of F, the formula C_x^0 is satisfiable, and then setting $x = \text{true}$ if $C_x^- = \varnothing$ ($x = \text{false}$ if $C_x^+ = \varnothing$) shows that F is satisfiable. Thus $<$ is a partial order. Let z be a minimal element in this order. Suppose $x \in V_F \backslash \{z\} \backslash V_z^+$. Then $C_x \subseteq C_z^-$ and we contradict the minimality of z. Thus $V_z^+ = V_F \backslash \{z\}$, and similarly $V_z^- = V_F \backslash \{z\}$. By Theorem 2 it follows that $|C_{F_z^+}| = |C_z^0| + |C_z^+| \geqslant |V_z^+| + 1 = |V_F|$, and similarly $|C_z^0| + |C_z^-| \geqslant |V_F|$. Since $|C_F| = |C_z^0| + |C_z^+| + |C_z^-| = |V_F| + 1$ this implies that $|C_z^+| = |C_z^-| - 1$ (we have already shown that $C_z^+ = \varnothing$ or $C_z^- = \varnothing$ is impossible).

Let c_1 be the single clause in which z appears positively and let c_2 be the clause in which z appears negatively. Let d_1, d_2 be such that $c_1 = d_1 \vee z$, $c_2 = d_2 \vee \bar{z}$. We show that $d_1 = d_2$. Suppose it is not the case. Then some variable y appears (say) positively in (say) d_1 and does not appears positively in c_2. Replace d_2 in F by $d_2 \vee y$ (if y apears negatively in d_2 this is equivalent to deleting d_2.) Since F is strongly minimal unsatisfiable there exists an assignment of truth values which satisfies the resulting formula. Clearly in this assignment $y = \text{true}$ and $z = \text{true}$, or else F itself would be satisfiable. But then changing the value of z to "false" would satisfy all clauses in F, a contradiction. We have thus shown that $F_z^+ = F_z^-$.

We now show that F_z^+ is strongly minimal unsatisfiable. Suppose that the formula H obtained by replacing some clause g in F_z^+ by $g \vee v$ is not satisfiable. If $g = d_1$ then replacing c_1 by $c_1 \vee v$ in F does not give a satisfiable formula: an assignment of truth values satisfying the resulting formula must have $v = \text{true}$, and then all clauses in H are satisfied. If $g \in C_z^0$ then replacing g by $g \vee v$ in F does not give a satisfiable formulai for, if an assignment of truth values satisfies the resulting formula then, since both c_1 and c_2 are satisfied, some variable other than z causes one of them to be satisfied, hence all clauses of Theorem 4 may be satisfied.

Since the number of variables in F_z^+ is one less than in F it follows by an induction hypothesis that the theorem holds for F_z^+ (note that when $|V_F| = 1$ the theorem holds trivially). Thus there exists a variable x appearing in all clauses of F_z^+, and since $F_z^+ = F_z^-$ it appears in all clauses of F, which proves (a). As before, F_x^+ and F_x^- are both minimal unsatisfiable, and hence $|C_{F_x^+}| = |C_x^+| \geqslant |V_x^+| + 1$ and similarly $|C_x^-| \geqslant |V_x^-| + 1$. Writing

$$\begin{aligned} |V_F| + 1 = |C_F| &= |C_x^+| + |C_x^-| \geqslant |V_x^+| + 1 + |C_x^-| + 1 \\ &\geqslant (|V_x^+| + |V_x^-| - |V_x^+ \cap V_x^-|) + 2 = |V_F| + 1 \end{aligned}$$

we deduce that equalities hold throughout, and thus $V_x^+ \cap V_x^- = \varnothing$, proving (b). Also, $|C_x^+| = |V_x^+| + 1$ and $|C_x^-| = |V_x^-| + 1$, proving (c).

Part (c) means that F_x^+ and F_x^- satisfy the same conditions as F, and hence the theorem can be applied to each of them, and recursively we descend until we reach formulas with one variable, which are of the form $y \wedge \bar{y}$. The theorem gives a prescription how to construct formulas fulfilling its conditions: take a variable x, split the rest of the variables into two disjoint sets, those variables appearing with x and those appearing with $\bar{x}$, in each set choose one "splitting" variable, and so on. A corollary of this observation is:

COROLLARY 4a. *If F is as in Theorem 4 then for each pair of clauses there exists a variable appearing positively in one and negatively in the other.*

The results of this section have a geometric interpretation. Let F be an unsatisfiable CNF formula in n variables. Let K be the cube $-1 \leqslant x_i \leqslant 1$, $i=1,\ldots,n$ in R^n. With every clause c of the form $c = \bigvee\{x_\alpha : \alpha \in A\} \vee \bigvee\{\bar{x}_\beta : \beta \in B\}$ we associate a box B_c contained in K, defined as

$$B_c = \{(x_1,\ldots,x_n) \in K : x_\alpha \geqslant 0 \quad \text{for} \quad \alpha \in A, x_\beta \leqslant 0 \quad \text{for} \quad \beta \in B\}.$$

Let $A_F = \{B_c : c \in C_F\}$. The "cell" $\{(x_1,\ldots,x_n) \in K : (-1)^{k_i} x_i \geqslant 0,\ i=1,\ldots,n$ and $k_i = 0$ or $k_i = 1$ for each $i\}$ is contained in B_c if and only if the assignment $x_i =$ true if $k_i = 1$, $x_i =$ false if $k_i = 0$ does not satisfy c. Thus F is unsatisfiable if and only A_F forms a cover of K. Minimal unsatisfiability of F corresponds to A_F being a minimal cover (i.e., no box can be deleted from it while keeping it a cover). Strong minimality of F means that whenever a box in A_F is halved by a hyperplane $x_i = 0$ and one half is deleted A_F ceases to be a cover. Finally, $V_F = \{x_1,\ldots,x_n\}$ means that every hyperplane $x_i = 0$ has a box supported by it. Theorem 4(a) says then that a cover A of K by $n+1$ boxes satisfying the above conditions has a hyperplane supporting all boxes. Corollary 4a says that such a cover is, in fact, a decomposition (i.e., no two boxes overlap). Parts (b) and (c) of Theorem 4 can be used to construct effectively all such covers, inductively.

We believe, but are unable to prove, that Theorem 4 holds also for infinite formulas. The condition $|C_F| = |V_F| + 1$ should be replaced by "there exists a matching from V_F into C_F in Γ_F, covering all elements of C_F but one." Part (c) of the theorem should be changed in a similar manner.

Another problem related to Theorem 4 is that of characterizing the finite minimal non-2-colorable hypergraphs $H = (V, E)$ for which $|V| = |E|$. Here too, in order to hope for a reasonable answer we have to assume strong minimality, which means that adding any new vertex to any edge makes H 2-colorable. But even in this case there are quite complicated examples. For example, the Fano plane has the above properties. Woodall describes in [6] a family of such hypergraphs.

Note added in proof. In his Ph.D. thesis Kassem [3] investigated questions closely related to the subject of this chapter. He considered decompositions of the cube in R^n denoted here by K into cells, which satisfy a condition he named "being neighborly." This means that the intersection of every two cells has dimension $n-1$. He obtained several characterizations for such decompositions.

References

1. R. Aharoni, Konig's duality theorem for infinite bipartite graphs, *J. London Math. Soc.* **29** (1984), 1–12.
2. R. Aharoni and P. Komjáth, On k-transversals, submitted.
3. G. Kassem, "Neighborly Decompositions of the n-Dimensional Cube," P.D. thesis, The Hebrew University, 1985.
4. N. Linial and M. Tarsi, Deciding hypergraph 2-colorability by H-resolution, *Theoret. Comput. Sci.* **38** (1985), 343–347.
5. L. Lovász, A generalisation of Konig's theorem, *Acta Math. Acad. Sci. Hungar.* **21** (1970), 443–446.
6. P. D. Seymour, On the two-colouring of hypergraphs, *Quart. J. Math. Oxford* **25** (1974), 303–312.
7. M. Tarsi, private communication.
8. D. R. Woodall, Property B and the four colour problem, in "Proc. Oxford Conf. On Combinatorial Math.," Southend-on-Sea, 1972, pp. 322–340.