A HOMOTOPY THEOREM FOR MATROIDS, I

вч

W. T. TUTTE

1. Introduction. By a matroid on a finite set M we understand a class M of non-null subsets of M which satisfies the following axioms.

AXIOM I. No member of M contains another as a proper subset.

AXIOM II. If $(X, Y) \in M$, $a \in X \cap Y$ and $b \in X - (X \cap Y)$, then there exists $Z \in M$ such that $b \in Z \subseteq (X \cup Y) - \{a\}$.

Such systems were introduced by Hassler Whitney [1].

As an example let L be any class of subsets of M forming a group under mod 2 addition, and let M be the class of all minimal non-null members of L. Then it is easily verified that L satisfies Axiom II and that each non-null member of L is a sum of non-null members of M. It follows that M satisfies both axioms and is thus a matroid. Such a matroid we call binary.

In particular M may be the set of edges of a finite graph G and L may be the class of 1-cycles mod 2 of G. Then it is found that the members of M are those sets of edges of G which define circuits. In this case we call M the circuit-matroid of G.

Given a matroid M let Q be the class of all unions of members of M. Then each element of Q is a subset of M. We partition Q into disjoint classes Q_{-1} , Q_0 , Q_1 , Q_2 , \cdots according to the following rules.

- (i) The null subset \emptyset of M, considered as an empty union, is the only member of Q_{-1} .
- (ii) When Q_r has been determined for $-1 \le r \le k$ we define Q_{k+1} as the class of all minimal members of

$$P_k = Q - \bigcup_{r=-1}^k Q_r.$$

That is Q_{k+1} consists of all members of P_k which have no other members of P_k as a subset. The members of Q are the flats of M. Those belonging to Q_d are the flats of dimension d, or d-flats.

At the end of §2 of this paper we interpret the dimensions of the flats of a circuit-matroid in terms of graph theory.

We shall see that the flats of a matroid M on a set M have some properties resembling those of the elements of a projective geometry. Because of this analogy we refer to the 0-flats, 1-flats and 2-flats of M as its points, lines and planes respectively. The points are simply the members of the class M.

We have to recognize one distinction which has no analogue in projective

Received by the editors October 26, 1956.

geometry. A flat F is disconnected if it can be represented as the union of two disjoint non-null subsets F' and F'' of M such that each point X of M satisfying $X \subseteq F$ satisfies also either $X \subseteq F'$ or $X \subseteq F''$. If no such representation is possible then F is connected. Thus the points of M and its (-1)-flat are connected.

A path in M is a finite sequence $P = (X_1, \dots, X_k)$ of one or more points of M, not necessarily all distinct, such that any two consecutive terms are distinct points of M which are subsets of the same connected line. The first and last terms of P are its origin and terminus respectively. If they are the same point we call P re-entrant. If P has only one term we call it degenerate.

If $P = (X_1, \dots, X_k)$ and $P' = (X_k, \dots, X_m)$ are paths of M such that the origin of P' is the terminus of P then we define their *product* PP' as the path $(X_1, \dots, X_k, \dots, X_m)$. Multiplication of paths is clearly associative. It is therefore permissible to write a path (PQ)R or P(QR) simply as PQR.

Suppose we have two paths PR and PQR where Q is either (i) of the form (X, Y, X) or (ii) of the form (X, Y, Z, X) with X, Y and Z subsets of the same plane. Then we say that each of PR and PQR can be derived from the other by an elementary deformation. Two paths P_1 and P_2 are homotopic if they are identical or if one can be derived from the other by a finite sequence of elementary deformations. Homotopy is clearly an equivalence relation. A path homotopic to a degenerate path is said to be null-homotopic.

In this paper we show that every re-entrant path in a matroid is null-homotopic. Actually we prove a more general theorem, as the result just stated is not sufficient for the purposes of Paper II. We first agree to call a subclass C of M convex if it has the following property: if two distinct members X and Y of C are subsets of the same line L then every point of M which is a subset of L is a member of C. Given a convex subclass C of C we say that a path C is off C if no term of C is a point of C. We then enquire into the condition that a path C off C can be transformed into a degenerate path by a finite sequence of elementary deformations so that all the intermediate paths are off C. In this paper we show how the idea of an elementary deformation must be generalized so as to make this transformation possible for every re-entrant path C off C.

It is hoped that the technique here developed for the study of matroids will be found useful in graph theory when applied to the circuit-matroids of graphs.

- 2. **Flats.** Let M be a matroid on a set M. We refer to the elements of M as the *cells* of the matroid. If S and T are subsets of M we use the symbol $S \subset T$ to denote that S is a proper subset of T. We write $\langle S \rangle$ for the union of all the points of M which are subsets of S. If S is a flat of M we denote its dimension by dS.
- (2.1) If S is a flat of M and k is an integer satisfying $-1 \le k < dS$ then there exists a flat T of M such that dT = k and $T \subseteq S$.

Proof. If the theorem fails let k be the greatest integer satisfying -1

 $\leq k < dS$ such that no k-flat T of **M** satisfies $T \subset S$. Clearly k > -1 and therefore $dS \geq 1$.

By the definition of k there exists a (k+1)-flat T' of M such that $T' \subseteq S$. By the definition of the classes Q_r we have

$$T' \in Q - \bigcup_{r=-1}^k Q_r \subset Q - \bigcup_{r=-1}^{k-1} Q_r = P_{k-1}.$$

But T' is not a minimal member of P_{k-1} , since dT' is not k. Hence there exists a minimal member T of P_{k-1} such that $T \subset T'$. But then dT = k. Since $T' \subseteq S$ this contradicts the definition of k. The theorem follows.

(2.2) If S and T are flats of **M** such that $S \subset T$, then dS < dT.

Proof. Since T is non-null we have dT > -1. If $dS \ge dT$ there is a flat U of M such that $U \subseteq S$ and dU = dT, by (2.1). But then T is not a minimal member of Q_{dT} , contrary to the definition of this class.

It follows from (2.2) that $\langle M \rangle$ has a greater dimension than any other flat of M.

It is convenient to say that a flat S is on a flat T if either $S \subseteq T$ or $T \subseteq S$. If S and T are distinct we can distinguish between the two cases by comparing dimensions.

(2.3) If S is a flat of M and $a \in S$, then $d(S - \{a\}) = dS - 1$.

Proof. If possible choose S and a so that $d\langle S - \{a\} \rangle \neq dS - 1$ and so that dS has the least value consistent with this.

By (2.1) there is a flat T of M such that $T \subseteq S$ and dT = dS - 1. Choose $b \in S - T$. Then $T \subseteq \langle S - \{b\} \rangle \subset S$. Hence $d\langle S - \{b\} \rangle = dS - 1$, by (2.2).

Suppose $a \in \langle S - \{b\} \rangle$. Then $\langle S - \{b\} \rangle \subseteq \langle S - \{a\} \rangle \subset S$. Hence $d\langle S - \{a\} \rangle = dS - 1$, by (2.2). But this is contrary to the choice of S and a.

We deduce that $a \in \langle S - \{b\} \rangle$. Hence there exists $X \in M$ such that $X \subseteq S$, $a \in X$ and $b \in X$. Since $b \in S$ there exists $Y \in M$ such that $Y \subseteq S$ and $b \in Y$. It follows by Axiom II that there exists $Z \in M$ such that $Z \subseteq S$, $a \in Z$ and $b \in Z$. $(Z = Y \text{ if } a \in Y)$. These results imply

$$(2.3a) \qquad \langle \langle S - \{a\} \rangle - \{b\} \rangle \subset \langle S - \{a\} \rangle \subset S,$$

$$(2.3b) \qquad \langle \langle S - \{b\} \rangle - \{a\} \rangle \subset \langle S - \{b\} \rangle \subset S.$$

We also have

$$(2.3c) \qquad \langle \langle S - \{a\} \rangle - \{b\} \rangle = \langle \langle S - \{b\} \rangle - \{a\} \rangle,$$

since each side of this equation represents the union of those points of M which include neither a nor b.

Since $d\langle S - \{b\} \rangle = dS - 1$ it follows from (2.3b) and the choice of S and a that $d\langle \langle S - \{b\} \rangle - \{a\} \rangle = dS - 2$. Hence $d\langle S - \{a\} \rangle = dS - 1$, by (2.2), (2.3a) and (2.3c). This contradiction establishes the theorem.

(2.4) Let S and T be flats of **M** such that $S \subseteq T$. Then there exists a flat U of **M** such that $U \subseteq T$, $\langle U \cap S \rangle = \emptyset$ and dU = dT - dS - 1.

Proof. Write $S_0 = S$, $T_0 = T$. If possible choose $a_0 \subseteq S_0$ and write $S_1 = \langle S_0 - \{a_0\} \rangle$, $T_1 = \langle T_0 - \{a_0\} \rangle$. Observe that $S_1 \subseteq T_1$. If possible choose $a_1 \in S_1$ and write $S_2 = \langle S_1 - \{a_1\} \rangle$, $T_2 = \langle T_1 - \{a_1\} \rangle$. Then $S_2 \subseteq T_2$. Continue this process until it terminates. By (2.3) this will be with S_k and T_k , where k = dS + 1 and $S_k = \emptyset$. Applying (2.3) to the sequence of the T_i we find that $dT_k = dT - k = dT - dS - 1$. We note that $\langle T_k \cap S \rangle \subseteq \langle S - \{a_0, \dots, a_{k-1}\} \rangle = S_k = \emptyset$. Hence the theorem is satisfied with $U = T_k$.

(2.5) If S and T are any flats of **M** then $d(S \cup T) + d(S \cap T) \ge dS + dT$.

Proof. Write $S_0 = S$. If possible choose $a_0 \in S_0 - (S_0 \cap T)$ and write $S_1 = \langle S_0 - \{a_0\} \rangle$. If possible choose $a_1 \in S_1 - (S_1 \cap T)$ and write $S_2 = \langle S_1 - \{a_1\} \rangle$, and so on. By (2.3) the process terminates with $S_k = \langle S \cap T \rangle$, where $k = dS - d\langle S \cap T \rangle$. We now have $T = S_k \cup T \subset S_{k-1} \cup T \subset \cdots \subset S_0 \cup T = S \cup T$. Hence $d(S \cup T) - dT \ge k = dS - d\langle S \cap T \rangle$, by (2.2).

Many "geometrical" results can be deduced from (2.2) and (2.5). For example any two distinct lines L_1 and L_2 on a plane P have a unique common point. To prove this we first use (2.2) to show that $d\langle L_1 \cap L_2 \rangle < dL_1 = 1$ and $L_1 \cup L_2 = P$. Then $d\langle L_1 \cap L_2 \rangle \ge 0$, by (2.5). Hence $d\langle L_1 \cap L_2 \rangle = 0$ and $\langle L_1 \cap L_2 \rangle$ is a single point of M. We can prove in the same way that if P_1 and P_2 are distinct planes on the same 3-flat E of M, then $\langle P_1 \cap P_2 \rangle$ is a line on E. Similarly if P is a plane and E a line on the same 3-flat E, and E is not on E, then $\langle P \cap L_1 \rangle$ is a point on E.

Not all the axioms of projective geometry are valid for matroids. For example two points are not necessarily on a common line. In general matroids are like geometrical figures but not like complete geometries.

Suppose M is the circuit-matroid of a graph G. If $S \in Q$ we write $G \cdot S$ for the subgraph of G made up of the edges of S and their incident vertices. We see that the flats S of M correspond to those subgraphs $G \cdot S$ in which each edge belongs to some circuit of the subgraph. In virtue of (2.3) dS+1 is the least number of edges which must be removed from $G \cdot S$ in order to destroy all its circuits, that is dS+1 is the rank or first Betti number of $G \cdot S$. The subgraph $G \cdot S$ is nonseparable if and only if the flat S is connected.

- 3. Connected flats. We begin this section with a study of the line.
- (3.1) Any line L of M is on at least two points. If X and Y are distinct points on L then $L = X \cup Y$. Moreover $X \cap Y$ is non-null if and only if L is connected.

Proof. Choose $a \in L$. Then $\langle L - \{a\} \rangle$ is a point on L, by (2.3). Choose $b \in \langle L - \{a\} \rangle$. Then $\langle L - \{b\} \rangle$ is a point on L, by (2.3), which is distinct from $\langle L - \{a\} \rangle$.

Let X and Y be distinct points on L. Then $X \subset X \cup Y \subseteq L$. Hence $X \cup Y = L$, by (2.2). If $X \cap Y$ is non-null then L is clearly connected. If $X \cap Y$ is null then either L is disconnected or there exists $Z \in M$ such that $Z \subseteq X \cup Y$ and Z meets both X and Y. In the latter case $X \subset X \cup Z \subset X \cup Y = L$, by Axiom I. This is impossible, by (2.2).

- (3.2) A disconnected line is on just two points, and a connected line is on at least three points.
- **Proof.** By (3.1) any two distinct points on a disconnected line L are disjoint and have L as their union. Hence L has at most two points, and therefore just two by (3.1).
- By (3.1) any connected line L has two distinct points X and Y, and we can find $a \in X \cap Y$. By (2.3) $\langle L \{a\} \rangle$ is a point on L distinct from X and Y. We shall need the following general theorems on connected flats.
- (3.3) Let S and T be connected flats of M such that $S \subset T$. Then there exists a connected (dS+1)-flat U of M which is on both S and T.
- **Proof.** Since T is connected we can find $X \in M$ such that $X \subseteq T$ and X meets both S and T-S. Choose such an X so that $S \cup X$ has the least possible number of cells. Clearly $S \cup X$ is a connected flat of M. Its dimension exceeds dS, by (2.2).
- Suppose $d(S \cup X) > dS + 1$. Choose $a \in (S \cup X) S$. Then $d((S \cup X) \{a\}) \ge dS + 1$, by (2.3). Hence there exists $Y \in M$ such that $Y \subseteq (S \cup X) \{a\}$ and Y meets $(S \cup X) S$. But $Y \cap S$ is null, by the choice of X. Hence $Y \subset X$, which is impossible by Axiom I. We deduce that $d(S \cup X) = dS + 1$. Hence the theorem is true with $U = S \cup X$.
- (3.4) Let S be a connected d-flat on a connected (d+2)-flat T of **M**. Then there exist distinct connected (d+1)-flats U and V of **M** such that $S = \langle U \cap V \rangle$ and $T = U \cup V$.
- **Proof.** By (3.3) there is a connected (d+1)-flat U which is on both S and T. Choose $a \in U S$ and write $W = \langle T \{a\} \rangle$. By (2.3) W is another (d+1)-flat on S and T. By (2.4) there is a line L on T having no point in common with S. It meets U and W in points X and Z respectively, by (2.5). (See Figure I.) By (2.2) we have $S \cup X = U$ and $S \cup Z = W$. Hence Z is not on U and therefore $U \cup Z = T$, by (2.2).

Assume W is not connected. Then $S \cap Z = \emptyset$.

Suppose $U \cap Z = \emptyset$. By the connection of T there exists $Z' \subseteq M$ such that $Z' \subseteq T$ and Z' meets both U and Z. Then $U \subset U \cup Z' \subset U \cup Z = T$, by Axiom I. This is impossible by (2.2). We deduce that $U \cap Z \neq \emptyset$. A similar argument in which X, S and U replace Z, U and T respectively shows that $X \cap S \neq \emptyset$. Choose $b \in Z \cap U$ and $c \in X \cap S$.

Write $V = \langle T - \{b\} \rangle$. By (2.3) V is a (d+1)-flat. It is on S since $b \in Z$ and $S \cap Z = \emptyset$. By (2.5) it has a common point Y with L. Clearly V is distinct from U and W and therefore Y is distinct from X and X, since $Y = S \cup Y$ by (2.2).

Now $c \in S \cap X \subseteq S \cap L = S \cap (Y \cup Z) = S \cap Y$, by (3.1). Hence V is connected.

If instead W is connected we write V = W.

We now have two connected (d+1)-flats U and V of M each of which is on both S and T. Hence $S \subseteq \langle U \cap V \rangle \subset U \subset U \cup V \subseteq T$, since U and V are

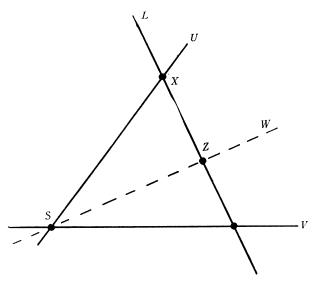


Fig. I

distinct and, by (2.2), neither is a subset of the other. In view of (2.2) this is possible only if $S = \langle U \cap V \rangle$ and $T = U \cup V$.

(3.5) Let S, T and U be flats of **M** such that S and T are connected, $S \cup U \subseteq T$ and $\langle S \cap U \rangle = \emptyset$. Then there exists a connected flat R of **M** such that $S \subseteq R \subseteq T$, $\langle R \cap U \rangle = \emptyset$ and dR = dT - dU - 1.

Proof. If possible choose S, T and U so that the theorem fails and dU has the least value consistent with this. Then dU > -1 since otherwise the theorem holds with T = R. Let W be a connected flat of M of greatest possible dimension such that $S \subseteq W \subseteq T$ and W does not contain U. Then dW = dT - 1 since otherwise, by (3.3) and (3.4), there exist distinct connected (dW+1)-flats K and L of M on T such that $\langle K \cap L \rangle = W$, and these cannot both contain U. By the choice of S, T and U there is a connected flat R of M such that $S \subseteq R \subseteq W \subset T$, $\langle R \cap U \rangle = \emptyset$ and $dR = dW - d\langle U \cap W \rangle - 1$. But then $dR \ge dT - dU - 1$, by (2.2) and (2.5). This contradiction establishes the theorem.

The foregoing results can be applied to circuit-matroids to obtain rather simple theorems about graphs. Thus from (3.5) with $S = \emptyset$ we find that if a nonseparable graph G has rank r and a subgraph $G \cdot U$ has a rank s then there is a nonseparable subgraph $G \cdot R$ of G of rank r-s having no circuit in common with $G \cdot U$.

- 4. The disconnected line. By a separation $\{S_1, S_2\}$ of a disconnected flat S of M we mean a pair of complementary non-null subsets of S such that each $X \subseteq M$ satisfying $X \subseteq S$ satisfies either $X \subseteq S_1$ or $X \subseteq S_2$.
 - (4.1) If $\{S_1, S_2\}$ is a separation of a flat S of M and X_1 and X_2 are

points of M such that $X_1 \subseteq S_1$ and $X_2 \subseteq S_2$, then $X_1 \cup X_2$ is a disconnected line of M.

- **Proof.** Suppose Y is a point on $X_1 \cup X_2$ distinct from X_1 and X_2 . Since $Y \subseteq S$ we have $Y \subseteq S_1$ or $Y \subseteq S_2$. Hence $Y \subset X_1$ or $Y \subset X_2$, contrary to Axiom I. Thus the only subsets of $X_1 \cup X_2$ which are flats of M are \emptyset , X_1 , X_2 and $X_1 \cup X_2$. Hence, by the definition of dimension, $X_1 \cup X_2$ is a line of M having $\{X_1, X_2\}$ as a separation.
- (4.2) Let L be a disconnected line on a connected d-flat S of M, where dS > 1. Then there exists a connected plane P of M such that $L \subset P \subseteq S$.
- **Proof.** Let the two points on L be X and Y. Let P be a connected flat of M of least possible dimension such that $L \subset P \subseteq S$. Assume dP > 2.

Suppose first that there is a disconnected line L', distinct from L, on X and P. Let its point other than X be Z. By (3.5) there is a connected (dP-2)-flat U on Y and P having no point in common with L'. By (3.4) there are distinct connected (dP-1)-flats V and W of M on P such that $\langle V \cap W \rangle = U$. By (2.2) and (2.5) V and W meet L' in distinct points. Since there are only two points on L' we may suppose X is on V. But then L is on V and the definition of P is contradicted. A similar argument applies if there is a disconnected line distinct from L on Y and P.

In the remaining case we choose $a \in P-L$ and write $R = \langle P - \{a\} \rangle$. Then $L \subseteq R$. Moreover dR = dP - 1, by (2.3). By the definition of P the flat R is disconnected. But there is no disconnected line on R, other than L, which is on either X or Y. Hence, by (4.1), the only possible separation of R is $\{X, Y\}$. Accordingly R = L and dP = 2, contrary to assumption. From this contradiction we deduce that P is a plane.

- (4.3) Let L be a disconnected line on a connected plane P of M. Let X and Y be the two points of L and let Z be any other point on P. Then $X \cup Z$ and $Y \cup Z$ are connected lines. Moreover they are the only lines of M which are on both Z and P.
- **Proof.** Any line on Z and P has a common point with the line $X \cup Y$. Hence, by (3.1), the only flats on Z and P which can be lines are $X \cup Z$ and $Y \cup Z$. By (3.4) both these flats must be connected lines.
- (4.4) Let L be a disconnected line on a connected plane P of M. Then every line on P other than L is connected.

Proof. Let L' be any such line. By (3.1) it is on a point Z distinct from X and Y. Hence, by (4.3) it is one of the connected lines $X \cup Z$ and $Y \cup Z$.

5. Convex subclasses. Convex subclasses of a matroid M were defined in the Introduction. As an example we may take the class of all points of M on a given flat. The convexity of this class follows from (3.1).

Consider any path $P = (X_1, X_2, \dots, X_k)$ of M. We say P is a path from X_1 to X_k . Any two consecutive terms of P have a non-null intersection, by (3.1). Hence the flat $X_1 \cup X_2 \cup \dots \cup X_k$ is connected. We denote this flat by F(P). If S is any flat of M such that $F(P) \subseteq S$ we say that P is a path on S.

(5.1) Let C be any convex subclass of M. Let S be a non-null connected flat

of M and let X and Y be points on S such that $Y \in \mathbb{C}$. Then there exists a path P from X to Y on S such that no term of P other than the first is a point of \mathbb{C} .

Proof. If possible choose S, X and Y so that the theorem fails and dS has the least value consistent with this. Clearly dS > 1. By (3.3) and (3.4) there is a connected (dS-2)-flat U and two distinct connected (dS-1)-flats V and W on S such that $X \subseteq U = \langle V \cap W \rangle$. Now Y is not on V or W, for otherwise there would be a path from X to Y on V or W of the kind required. By (3.5) there is a connected line L on S and Y such that $\langle L \cap U \rangle = \emptyset$. This meets V and W in distinct points Z(V) and Z(W) respectively, by (2.2) and (2.5). At least one of these, say Z(V), belongs to M - C since C is convex. By the choice of S, X and Y there is a path Q from X to Z(V) on V such that no term of Q other than the first is a point of C. Adjoining Y to Q we obtain a path P from X to Y on S of the kind required. This contradiction establishes the theorem.

We now distinguish four kinds of re-entrant paths of M as elementary with respect to a given convex subclass C of M. The first kind consists of all paths off C of the form (X, Y, X). The second consists of all paths off C of the form (X, Y, Z, X) such that $d(X \cup Y \cup Z) \leq 2$.

Suppose P is a plane of M on which there are two distinct points A and B of C such that each connected line on P is on either A or B. Then any path off C on P of the form (X, Y, Z, T, X) such that X, Y, Z and T are distinct, the lines $X \cup Y$ and $Z \cup T$ are on A, and the lines $Y \cup Z$ and $T \cup X$ are on B is an elementary re-entrant path of the third kind with respect to C.

Suppose E is a 3-flat of M on which there are three points A, B and C such that $A \cup B$, $B \cup C$ and $C \cup A$ are disconnected lines. Let there be just six connected planes on E, two on each of these disconnected lines. Suppose A, B and C are all in M - C but there are two distinct members of C on each of the six connected planes. Then any path off C of the form (A, X, B, Y, A), where X and Y are on distinct connected planes on $A \cup B$ and E, is an elementary re-entrant path of the fourth kind with respect to C.

In studying the preceding case it is convenient to use the following notation. We write Z_1 , Z_2 and Z_3 for A, B and C. We enumerate the six connected planes as P_1, \dots, P_6 in such a way that $\langle P_i \cap P_{i+3} \rangle = Z_j \cup Z_k$, where $1 \leq i \leq 3$ and (i, j, k) is a permutation of (1, 2, 3). In general we write $\langle P_i \cap P_j \rangle = L_{ij}$ for $1 \leq i < j \leq 6$. If j = i + 3 then L_{ij} is the disconnected line $(Z_1 \cup Z_2 \cup Z_3) - Z_i$. If $j \neq i + 3$ let k be that integer 1, 2 or 3 which is not congruent to i or j mod 3. Then L_{ij} is on Z_k and it meets P_k and P_{k+3} in two distinct points. It is therefore connected, by (3.2). Clearly it is on no connected plane on E other than P_i and P_j . The 12 lines $L_{ij}, j \neq i + 3$, are the only connected lines on E, for by (3.4) any connected line on E is on two distinct connected planes on E.

We write $\langle P_i \cup P_j \cup P_k \rangle = X_{ijk}$ for $1 \le i < j < k \le 6$. Then X_{ijk} is a point of M, being identical with $\langle L_{ij} \cap P_k \rangle$. If two of the suffices i, j and k are congruent mod 3 then X_{ijk} is one of the points Z_1 , Z_2 and Z_3 . The remaining eight points X_{ijk} are all distinct, for on any one of them there can be only three

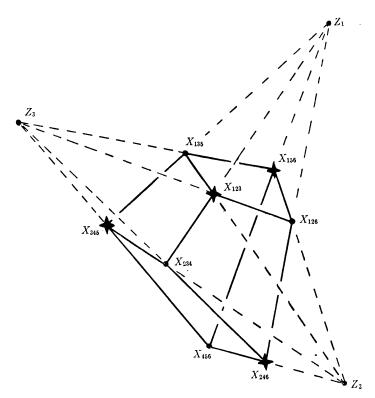


Fig. II

distinct planes such that each is on one of the lines $Z_1 \cup Z_2$, $Z_2 \cup Z_3$ and $Z_3 \cup Z_1$. These eight points, together with Z_1 , Z_2 and Z_3 , are the only points on E. For any point on E is on three distinct connected planes on E, by two applications of (3.4). (See Figure II.)

Consider the plane P_1 . The only points on it are Z_2 , Z_3 , X_{123} , X_{126} , X_{135} and X_{156} . We may adjust the notation so that $X_{123} \in \mathbb{C}$. The other point of \mathbb{C} on P_1 can have no common connected line with X_{123} and must therefore be X_{156} . We now find that $X_{246} \in \mathbb{C}$ since this is the only point on P_2 having no common connected line with X_{123} . Proceeding in this way we find that $X_{ijk} \in \mathbb{C}$ if and only if no two of the suffices are congruent mod 3 and the number of suffices less than 4 is odd. In Figures II, III and IV we represent points of \mathbb{C} by four-pointed stars.

To construct a matroid having the structure just described we may use a method based on (2.3). We take M to be a set of six cells in 1-1 correspondence with the planes P_i . Any point X_{ijk} is represented by the set of those cells not corresponding to planes on X_{ijk} .

Suppose we have two paths PQR and PR off C, where Q is an elementary re-entrant path of the kth kind with respect to C. Then we call the process

of deriving one of the paths PQR and PR from the other an elementary deformation of the kth kind with respect to C. We say that two given paths P' and P'' off C are homotopic with respect to C (written $P' \sim P''$ (C)) if they are identical or if one can be derived from the other by a finite sequence of elementary deformations with respect to C. Homotopy with respect to C is an equivalence relation. A path P homotopic to a degenerate path with respect to C is said to be null-homotopic with respect to C (written $P \sim O(C)$).

The null subset of M is clearly convex. If C is null we have only elementary deformations of the first and second kind to consider, and homotopy with respect to C becomes identical with the homotopy defined in the Introduction.

If P is any path of M we write P^{-1} for the path obtained by taking the terms of P in reverse order.

(5.2) If P is any path off C then $PP^{-1} \sim 0$ (C).

Proof. If possible choose P so that the theorem fails and P has the least number s of terms consistent with this. If s>1 we can write P=QR, where Q and R have each fewer than s terms. Since RR^{-1} and QQ^{-1} can be converted into degenerate paths by elementary deformations we have $PP^{-1}=QRR^{-1}Q^{-1}\sim QQ^{-1}\sim 0$ (C). If s=1 then PP^{-1} is an elementary re-entrant path of the first kind, and so $PP^{-1}\sim 0$ (C). The theorem follows.

(5.3) If PUR and PVR are paths off C such that $UV^{-1}\sim 0$ (C), then $PUR\sim PVR$ (C).

Proof. By (5.2) we have $V^{-1}V = V^{-1}(V^{-1})^{-1} \sim 0$ (C). Hence $PUR \sim PUV^{-1}VR \sim PVR$ (C).

(5.4) Let C be any convex subclass of a matroid M. Let S be a d-flat of M on a (d+1)-flat T of M. Suppose all the points on S and at least one other point on T are members of C. Then all the points on T are members of C.

Proof. Suppose the theorem false. Then we can find points $X \in \mathbb{C}$ and $Y \notin \mathbb{C}$, both on T but not on S. The flat $X \cup Y$ is connected since otherwise $S \subset S \cup X \subset T$, contrary to (2.2). By (5.1) there is a path from X to Y on $X \cup Y$ whose second term, X' say, is not a member of C. By (2.5) the line $X \cup X'$ has a point X'' in common with S. But $X'' \notin C$, by the definition of a convex subclass. This is contrary to hypothesis.

6. Proof of the main theorem.

(6.1) Let C be any convex subclass of a matroid M and let P be any re-entrant path of M off C. Then $P \sim 0$ (C).

Proof. Assume the theorem false. Let P be any re-entrant path off C which is not null-homotopic with respect to C, and for which dF(P) has the least value, n say, consistent with this condition. For an arbitrary path Q of M we call dF(Q) the dimension of Q.

By far the most difficult part of the proof is that covered by the following lemma.

LEMMA. Suppose $n \ge 3$. Let Q = (W, X, Y, Z, W) be a path off C of dimension

n such that $W \cup X \cup Y$ and $Y \cup Z \cup W$ are connected planes and $W \cup Y$ is a disconnected line. Then $Q \sim 0$ (C).

Proof. Write $F_1 = W \cup X \cup Y$ and $F_2 = Y \cup Z \cup W$.

We note that if Q' = (W, X', Y, Z', W) is a path off C such that X' is on F_1 and Y' on F_2 , then

$$(6.1a) Q' \sim Q(C).$$

For

$$Q' \sim (W, X', Y)(Y, X, W)(W, X, Y)(Y, Z, W)(W, Z, Y)(Y, Z', W)$$
 (C)

by (5.2). But (W, X', Y)(Y, X, W) and (W, Z, Y)(Y, Z', W) are re-entrant paths off C of dimension < n and are therefore null-homotopic with respect to C. Hence $Q' \sim (W, X, Y)(Y, Z, W) = Q(C)$.

A transversal of dimension n-1 is a connected (n-1)-flat of M which is on F(Q) but not on both W and Y. By (2.2) and (2.5) such a transversal meets each of F_1 and F_2 in a line. These two lines are connected, by (4.4).

A transversal of dimension n-2 is a connected (n-2)-flat of M which is on F(Q) but not on W or Y. By (2.2) and (2.5) the transversal has just one point in common with each of F_1 and F_2 . We call these two points the poles of the transversal.

Let B be any transversal of dimension n-2, with poles X' on F_1 and Z' on F_2 . Then B is on two distinct connected (n-1)-flats of M on F(Q), by (3.4). Using (2.5) we find that each of these is on one, but not both, of W and Y. Hence, by (2.2) they are $B \cup W$ and $B \cup Y$. They are transversals of dimension n-1. The flats $X' \cup W$, $X' \cup Y$, $Z' \cup W$ and $Z' \cup Y$ are their connected lines of intersection with F_1 and F_2 . We note that a path (W, X', Y, Z', W) exists.

Assume that Q is not null-homotopic with respect to C.

Suppose B is a transversal of dimension n-2 with poles X' on F_1 and Z' on F_2 . Suppose further that neither X' nor Z' belongs to C. Then, by (5.1) there is a path R off C from X' to Z' on B. Now (W, X')R(Z', W) and $(X', Y, Z')R^{-1}$ are paths on the (n-1)-flats $B \cup W$ and $B \cup Y$ respectively. Hence their dimensions are less than n and so they are null-homotopic with respect to C. Using (6.1a), (5.3) and (5.2) we find $Q \sim (W, X', Y, Z', W) = (W, X')(X', Y, Z')(Z', W) \sim (W, Z')R^{-1}R(Z', W) \sim 0$ (C). This is contrary to assumption. We deduce that each transversal of dimension n-2 has at least one pole in C.

By (3.5) there is a transversal A of dimension n-1 which is not on Y. Let its lines of intersection with F_1 and F_2 be L_1 and L_2 respectively. They are connected lines on W. By (3.2) there is a point X' of M-C other than W on L_1 . By (3.5) there is a connected (n-2)-flat B of M which is on A and X' but not on W. Now B is a transversal of dimension n-2. Let its pole on L_2 be U_2 . Then $U_2 \subset C$. Similarly there is a transversal B' of dimension n-2 on A having a point A' of A' of

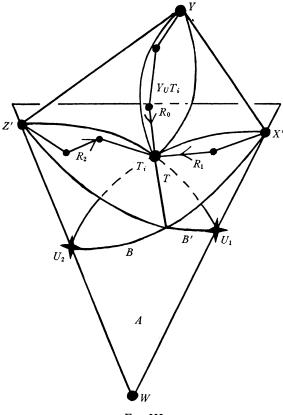


Fig. III

as its pole on L_1 . (See Figure III.) We write $T = \langle B \cap B' \rangle$. By (2.2) and (2.5) T is an (n-3)-flat of M.

Let **S** be the class of all members of M-C on T. Since $T \subset B$, $X' \in M-C$ and $U_2 \in C$ it follows by (5.4) that **S** is non-null.

Let T_i be any point of S. Suppose that the flat $Y \cup T_i$ is connected. Then there is a path R_0 from Y to T_i on $Y \cup T_i$ which is off C, by (5.1). Similarly there is a path R_1 from X' to T_i on B and a path R_2 from Z' to T_i on B', both R_1 and R_2 being off C. Now $(X', Y)R_0R_1^{-1}$ is a re-entrant path on the transversal $B \cup Y$ of dimension n-1, and $(Y, Z')R_2R_0^{-1}$ is a re-entrant path on the transversal $B' \cup Y$ of dimension n-1. Hence both these paths are null-homotopic with respect to C. Applying (6.1a), (5.3) and (5.2) we find $Q \sim (W, X', Y, Z', W) = (W, X')(X', Y)(Y, Z')(Z', W) \sim (W, X')R_1R_0^{-1}R_0R_2^{-1} \cdot (Z', W) \sim (W, X')R_1R_2^{-1}(Z', W)$ (C). But the last path is on the (n-1)-flat A and is therefore null-homotopic with respect to C. Hence $Q \sim 0$ (C), contrary to assumption. We deduce that $Y \cup T_i$ has a separation $\{Y, T_i\}$. That is $Y \cup T_i$ is a disconnected line, and Y and T_i are the two points on it.

We can repeat the above argument with $B \cup Y$ replacing A. Instead of

B' we then obtain a transversal B'' of dimension n-2 on $B \cup Y$ with its pole on $U_2 \cup Y$ a member of M-C and its pole on $X' \cup Y$ a member of C. We denote the (n-3)-flat $(B \cap B'')$ by T'. We find that any point T' of M-C on T' is such that $W \cup T'$ is a disconnected line. But, by (2.5), B'' has a point in common with the disconnected line $Y \cup T_i$, and this point can only be T_i . Hence T_i is one of the points of M-C on T'.

We conclude that any point T_i of **S** is such that $W \cup T_i$ and $Y \cup T_i$ are disconnected lines.

Let E be a connected flat of M on F_1 and F(Q) which is on some point of S and has the least dimension consistent with this property. It is clear that either F(Q) or one of its subsets satisfies these conditions. We have

$$(6.1b) n = dF(Q) \ge dE \ge 3,$$

$$(6.1c) d\langle E \cap T \rangle \ge dE - 3,$$

by (2.2) and (2.5). Choose a point N on $\langle E \cap T \rangle$, taking $N \in \mathbb{C}$ if this is possible. By (3.5) there is a connected (dE-1)-flat E' of M on F_1 and E but not on N. By (2.2) and (2.5) $\langle E' \cap T \rangle$ is a $(d\langle E \cap T \rangle - 1)$ -flat on $\langle E \cap T \rangle$. All the points of M which are subsets of $\langle E' \cap T \rangle$ belong to C, by the definition of E. By the choice of N this implies that either $d\langle E' \cap T \rangle = -1$ or $N \in C$. But in the latter case all the points on $\langle E \cap T \rangle$ belong to C, by (5.4), contrary to the definition of E. Hence $d\langle E' \cap T \rangle = -1$ and therefore $d\langle E \cap T \rangle = 0$. Hence dE = 3, by (6.1b) and (6.1c). Henceforth we use the symbol T_i to denote the single point $\langle E \cap T \rangle$, which must be in S.

Suppose $n \ge 4$. Then F_2 is not on E. By (3.5) there is a connected (n-1)-flat E'' of M on F_2 and F(Q) but not on T_i . Write $F_3 = \langle E'' \cap E \rangle$. Then F_3 is a plane on E and $W \cup Y$, by (2.2) and (2.5). By (3.5) there is a connected line E on E and E on E and E on E and E on E on E on the interval of the in

By (5.1) there is a path R from Y to W on F_3 which is off C. The re-entrant paths (W, X, Y)R and $(Y, Z, W)R^{-1}$ are on E and E'' respectively and so have dimensions < n. Hence they are null-homotopic with respect to C. Using (5.3) and (5.2) we find $Q = (W, X, Y)(Y, Z, W) \sim R^{-1}R \sim 0$ (C), contrary to assumption.

We deduce, using (6.1b), that n=3. This implies dT=0. Hence T is a point of M, identical with T_i and therefore a member of M-C. The three flats $W \cup Y$, $Y \cup T$ and $T \cup W$ are disconnected lines. The flat $W \cup Y \cup T$ is not a connected plane of M by (4.3).

Any plane on F(Q) has a point in common with each of the disconnected

lines $W \cup Y$, $Y \cup T$ and $T \cup W$. It is therefore on one of these lines. Each line of F(Q) is on a plane of F(Q), by (2.2) and (2.3). It follows that each line on F(Q) is on one of the points W, Y and T.

Let F be any transversal of dimension 2. It meets F_1 and F_2 in connected lines L_1 and L_2 respectively. Let the points on L_1 other than W or Y be X_1, \dots, X_k . By (3.4) there is a transversal B_i of dimension 1 on F and X_i for each i. The line B_i must be on T. Hence $B_i = X_i \cup T$, and B_i is uniquely determined for each i. Let X_i' denote the point of intersection of B_i and L_2 . Since $B_i = T \cup X_i'$ for each i the k points X_1', \dots, X_k' are all distinct. But at most one point on each of the lines L_1 and L_2 belongs to C, and no transversal of dimension 1 has both its poles in M - C. Applying (3.2) we deduce that k = 2. Moreover we can adjust the notation so that $(X_1, X_2') \subseteq C$ and $(X_2, X_1') \subseteq M - C$.

Distinct lines on T and F meet L_1 in distinct points, by (3.1). Hence the only connected lines on T and F are B_1 and B_2 . But each point of F is on two connected lines on F, and one of these is on T. Hence X_1 and X_2' are the only points of C on F. We thus prove that each connected plane on F(Q) not on $W \cup Y$ is on just two points of C.

Any connected line on F_1 is on a transversal of dimension 2, by (3.4). Hence it is on just three points, one of which is in C. Distinct lines on F_1 and W (or Y) meet a given connected line on F_1 and Y (or W) in distinct points, by (3.1). It follows that on F_1 there are just two connected lines on each of the points W and Y. As each point on F_1 is on two connected lines on F_1 we deduce that F_1 is on just two points of C. Analogous results hold for F_2 . Two distinct transversals of dimension 2 are both on T and therefore meet F_1 in distinct lines, by (2.2). Accordingly there are just two connected planes on F(Q) and $T \cup W$, and just two on F(Q) and $Y \cup T$ (since $W \cup Y \cup T$ is not a connected plane).

It follows from these results that either Q is an elementary re-entrant path of the fourth kind with respect to C or there is a third connected plane F_3 on F(Q) and $W \cup Y$. The first alternative must be rejected since it implies $Q \sim 0$ (C).

Let the points of intersection with F_3 of B_1 and B_2 be X_1'' and X_2'' respectively. These are both in M-C. Each is on two connected lines on F_3 , one on W and the other on Y, by (4.3). We have $(W, X, Y, X_1'', W) = (W, X, Y) \cdot (Y, X_1'', W) \sim 0$ (C). For otherwise we can repeat the first parts of the preceding proof with (W, X, Y, X_1'', W) replacing Q and obtain a contradiction, for the transversal B_2 then has both poles in M-C. Similarly $(Y, Z, W, X_1'', Y) = (Y, Z, W)(W, X_1'', Y) \sim 0$ (C). Applying (5.3) and (5.2) we find $Q = (W, X, Y)(Y, Z, W) \sim (W, X_1'', Y)(Y, X_1'', W) \sim 0$ (C), contrary to assumption. The lemma follows.

We return to the path P defined at the beginning of this proof. We note that $n=dF(P) \ge 1$ since otherwise P would be trivially null-homotopic with respect to C. We choose a connected (n-1)-flat E of M which is on F(P)

and the origin X_0 of P. This choice is possible, by (3.3).

Let $R = (X_0, \dots, X_m, X_0)$ be any re-entrant path with the same origin as P on F(P). We write u(R) for the number of terms of R, counting repetitions, which are not on E. If u(R) > 0 we write X_i for the first term of R which is not on E. We then write $v(R) = d(X_{i-1} \cup X_i \cup X_{i+1})$, taking $X_{m+1} = X_0$ if i = m. If u(R) = 0 we write v(R) = 0.

Henceforth we suppose R chosen so as to satisfy the following conditions:

- (i) $R \sim P(C)$,
- (ii) u(R) has the least value consistent with (i),
- (iii) v(R) has the least value consistent with (i) and (ii).

We consider first the case u(R) > 0. Then v(R) > 0. We may conveniently write R in the form $R_1(X_{i-1}, X_i, X_{i+1})R_2$, noting that R_1 is a path on . We write also $F = X_{i-1} \cup X_i \cup X_{i+1}$.

Suppose v(R) = 1. Then F is a connected line. If $X_{i+1} = X_{i-1}$ we have $R \sim R_1 R_2$ (C), by an elementary deformation of the first kind. This is impossible since $u(R_1 R_2) < u(R)$. If $X_{i+1} \neq X_{i-1}$ then $(X_{i-1}, X_i, X_{i+1}, X_{i-1})$ is an elementary re-entrant path of the second kind with respect to C. Applying (5.3) we find $R \sim R_1(X_{i-1}, X_{i+1})R_2$ (C). This is impossible since

$$u(R_1(X_{i-1}, X_{i+1})R_2) < u(R).$$

Suppose v(R) = 2. Then F is a connected plane on F(Q). It meets E in a line L, by (2.2) and (2.5). Let Z be the point of intersection of the lines L and $X_i \cup X_{i+1}$ on F. We discuss first the case $Z \in \mathbf{M} - \mathbf{C}$. In this case we define Q as the degenerate path (Z) if $Z = X_{i+1}$ and as the path (Z, X_{i+1}) otherwise. Then $(X_i, X_{i+1})Q^{-1}(Z, X_i)$ is an elementary re-entrant path of the first or second kind. Hence $R \sim R_1(X_{i-1}, X_i, Z)QR_2(\mathbf{C})$, by (5.3). If L is connected we have $(X_{i-1}, X_i, Z, X_{i-1}) \sim 0$ (C) and therefore $R \sim R_1(X_{i-1}, Z)QR_2(\mathbf{C})$, by (5.3). If L is not connected it is on a connected plane F' of M on E, by (4.2). We can find a connected line L' on X_{i-1} and F', and a point T of M - C distinct from X_{i-1} on L'. Then $T \cup X_{i-1}$ and $T \cup Z$ are connected lines, by (4.3). Using the lemma and the definition of n we find $(X_{i-1}, X_i, Z, T, X_{i-1}) \sim 0$ (C). Hence $R \sim R_1(X_{i-1}, T, Z)QR_2(C)$, by (5.3). So whether L is connected or not we have $R \sim R_3QR_2(C)$, where R_3 is on E. This is impossible since $u(R_3QR_2) < u(R)$.

We go on to the case $Z \in C$, illustrated in Figure IV. By (3.4) there is a connected line L' other than $X_i \cup X_{i+1}$ on X_{i+1} and F. If L' is on X_{i-1} we have $R \sim R_1(X_{i-1}, X_{i+1})R_2$ (C), using (5.3) with the elementary re-entrant path $(X_{i-1}, X_i, X_{i+1}, X_{i-1})$ of the second kind. This is impossible since $u(R_1(X_{i-1}, X_{i+1})R_2) < u(R)$. Hence L' must meet the lines $X_{i-1} \cup X_i$ and L in distinct points U and V respectively. Since $Z \in C$ and $X_{i-1} \in M - C$ we have $V \in M - C$.

Suppose $U \in M - C$. Using (5.3) with elementary re-entrant paths of the first and second kinds we find

 $R \sim R_1(X_{i-1}, U, X_i, X_{i+1})R_2 \sim R_1(X_{i-1}, V, U, X_{i+1})R_2 \sim R_1(X_{i-1}, V, X_{i+1})R_2(C).$ This is impossible since $u(R_1(X_{i-1}, V, X_{i+1})R) < u(R).$

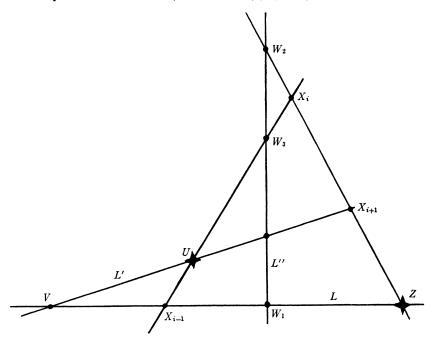


Fig. IV

Suppose $U \in \mathbb{C}$. It may happen that each connected line on F is on either U or Z. Then $(X_{i-1}, X_i, X_{i+1}, V, X_{i-1})$ is an elementary re-entrant path of the third kind with respect to C. Using (5.3) we have $R \sim R_1(X_{i-1}, V, X_{i+1})R_2$ (C), which is impossible, as before. Hence there is a connected line L'' on F which is not on U or Z. If L'' is on X_{i+1} we can substitute it for L' in the preceding argument and so reduce to the case $U \in M - C$. We may therefore suppose L'' is not on X_{i+1} .

If L'' is on X_i it meets L in a point W_1 distinct from X_{i-1} and Z. Writing $R' = R_1(X_{i-1}, W_1, X_i, X_{i+1})R_2$ we have $R' \sim R(C)$, by (5.3). If L'' is not on X_i it meets $X_i \cup X_{i+1}$ in a point W_2 distinct from X_i , X_{i+1} and Z. If L'' is then on X_{i-1} we write $R' = R_1(X_{i-1}, W_2, X_{i+1})R_2$ and have

$$R' \sim R_1(X_{i-1}, W_2, X_i, X_{i+1})R_2 \sim R(C),$$

by (5.3). If instead L'' is not on X_{i-1} it meets the lines L, $X_i \cup X_{i+1}$ and $X_{i-1} \cup X_i$ in distinct points W_1 , W_2 and W_3 respectively of M - C. We then write $R' = R_1(X_{i-1}, W_1, W_2, X_{i+1})R_2$ and have $R' \sim R_1(X_{i-1}, W_1, W_3, W_2, X_i, X_{i+1})R_2 \sim R_1(X_{i-1}, W_3, X_i, X_{i+1})R_2 \sim R$ (C), by (5.3). For each of these three possibilities we have $R' \sim R$ (C), u(R') = u(R) and v(R') = v(R) = 2. Hence we may

replace R by R' in the preceding argument. This reduces the problem to the case $U \in M - C$, which we have found to lead to a contradiction.

We now consider the case v(R) > 2. By (3.3) there is a connected plane K on $X_{i-1} \cup X_i$ and the connected v(R)-flat F. This plane meets E in a line L. Choose a point T distinct from X_{i-1} on L and if possible in C. By (3.5) there is a connected (v(R)-1)-flat F' on $X_i \cup X_{i+1}$ and F but not on T. Now F' is not on X_{i-1} , for otherwise we would have $F \subseteq F'$, contrary to (2.2). Hence F' meets L in a point T' distinct from X_{i-1} and T. It follows that L is connected, by (3.2), and that $T' \subseteq M - C$. The flats K and F' intersect in a line L' on X_i and T'. If L' is connected we write

$$R' = R_1(X_{i-1}, T', X_i, X_{i+1})R_2$$

and have $R' \sim R(C)$, by (5.3). If L' is not connected it is on a connected plane K' on F', by (4.2). K' meets E in a connected line L'' on T', by (4.4). We can find a point U on L'' distinct from T' and in M - C. The flat $U \cup X_i$ is a connected line, by (4.3). Using the lemma and the definition of n we find $(T', U, X_i, X_{i-1}, T') \sim 0$ (C). In this case we write

$$R' = R_1(X_{i-1}, T', U, X_i, X_{i+1})R_2.$$

Then by (5.2) and (5.3) we have

$$R' \sim R_1(X_{i-1}, T', U, X_i, X_{i-1}, T', X_{i-1}, X_i, X_{i+1})R_2$$

$$\sim R_1(X_{i-1}, T', X_{i-1}, X_i, X_{i+1})R_2 \sim R(C).$$

So whether L' is connected or not we have $R' \sim R$ (C), u(R') = u(R) and v(R') < v(R), which is contrary to the definition of R.

From the above analysis we deduce that u(R) = 0. Hence R is on E and has dimension $\langle n \rangle$. Hence $P \sim R \sim 0$ (C), contrary to assumption. The theorem follows.

7. Special cases. With C null in (6.1) we find that every re-entrant path in a matroid M is null-homotopic, as stated in the Introduction.

In applying this result to the circuit-matroid M of a graph G we must remember that a path in M corresponds to a sequence of circuits of G such that any two consecutive circuits form a nonseparable subgraph of rank 2. It can be shown that such a subgraph is made up of three arcs such that any two have both ends but no other edge or vertex in common. Each of the elementary deformations by which a re-entrant sequence of circuits can be transformed into a sequence with only one member operates within some non-separable subgraph of rank ≤ 3 .

REFERENCE

1. Hassler Whitney, The abstract properties of linear dependence, Amer. J. Math. vol. 57 (1935) pp. 507-533.

University of Toronto, Toronto, Ont.