# A COMBINATORIAL DECOMPOSITION THEORY 

WILLIAM H. CUNNINGHAM AND JACK EDMONDS

1. Introduction. Given a finite undirected graph $G$ and $A \subseteq E(G)$, $G(A)$ denotes the subgraph of $G$ having edge-set $A$ and having no isolated vertices. For a partition $\left\{E_{1}, E_{2}\right\}$ of $E(G), W\left(G ; E_{1}\right)$ denotes the set $V\left(G\left(E_{1}\right)\right) \cap V\left(G\left(E_{2}\right)\right)$. We say that $G$ is non-separable if it is connected and for every proper, non-empty subset $A$ of $E(G)$, we have $|W(G ; A)| \geqq$ 2. A split of a non-separable graph $G$ is a partition $\left\{E_{1}, E_{2}\right\}$ of $E(G)$ such that

$$
\left|E_{1}\right| \geqq 2 \leqq\left|E_{2}\right| \text { and }\left|W\left(G ; E_{1}\right)\right|=2 .
$$

Where $\left\{E_{1}, E_{2}\right\}$ is a split of $G, W\left(G ; E_{1}\right)=\{u, v\}$, and $e$ is an element not in $E(G)$, we form graphs $G_{i}, i=1$ and 2 , by adding $e$ to $G\left(E_{i}\right)$ as an edge joining $u$ to $v$. In this situation we write $G \rightarrow\left\{G_{1}, G_{2}\right\}$, and call $\left\{G_{1}, G_{2}\right\}$ a simple decomposition of $G$, associated with the split $\left\{E_{1}, E_{2}\right\}$ and the marker $e$. This paper describes a unique decomposition theory which includes among its applications a theory of graph decomposition based on this notion of simple decomposition. In this section we continue with the description of this instance of the theory.

Let $G$ be a non-separable graph. A decomposition $D$ of $G$ is defined inductively to be either $\{G\}$ or a set obtained from a decomposition $D^{\prime}$ of $G$ by replacing a member $G_{1}$ of $D^{\prime}$ by the members of a simple decomposition of $G_{1}$, where the marker of this simple decomposition is not an edge of any member of $D^{\prime}$. If $D^{\prime \prime}$ is obtained from $D$ by a (non-empty) sequence of operations of the kind described above, then $D^{\prime \prime}$ is said to be a (strict) refinement of $D$. If the sequence consists of exactly one operation, the refinement is simple.

We can associate a graph $T$ with any decomposition $D$ of a nonseparable graph $G$. The vertices of $T$ are the members of $D$ and the edges are the markers of $D$; each marker joins in $T$ the two members of $D$ of which it is an edge. It is clear that $T$ is a tree. This "decomposition tree" provides a useful way to visualize a decomposition.

Two decompositions $D, D^{\prime}$ of $G$ are equivalent if $D^{\prime}$ can be obtained from $D$ by replacing some of the markers of $D$ by markers of $D^{\prime}$. All unique decomposition theorems of this paper involve uniqueness "up to equivalence", but we will tend not to include this phrase in their statements. The decomposition $D$ of $G$ is minimal with some property $P$ if $D$ has $P$ and there does not exist a decomposition $D^{\prime}$ of $G$ also having $P$,

Received September 19, 1978 and in revised form May 2, 1979.
such that $D$ is a strict refinement of $D^{\prime}$. A decomposition $D$ is trivial if $|D|=1$. A non-separable graph $G$ is prime if it has no non-trivial decomposition. We observe that the prime graphs are precisely those which are 3 -connected in the sense of Tutte [23].

One might hope that each non-separable graph $G$ has a unique decomposition consisting of prime graphs. Examples of graphs which are badly behaved in this regard are the polygons (connected graphs in which each vertex has degree 2) and bonds (connected graphs having 2 vertices and no loops). Bonds and polygons having 4 or more edges have at least two different (inequivalent) prime decompositions. (A bond or polygon having 6 or more edges can have two prime decompositions having nonisomorphic decomposition trees.) Other types of graphs can also have more than one prime decomposition, but bonds and polygons play a special role in the uniqueness theory. The following result is the main unique decomposition theorem for graphs; its proof (as well as a discussion of its relation to other work on graph decomposition) appears in Section 4.

Theorem 1. Let $G$ be a non-separable graph. Then $G$ has a unique minimal decomposition, each of whose members is prime, a polygon, or a bond.

In the remainder of this section, we derive some elementary properties of simple decompositions and splits of non-separable graphs. These properties will motivate the definition of "decomposition frame" (Section 2 ), which is the context of the main theorems. It is convenient at this point to introduce some notation. If $G$ is a non-separable graph, $\left\{E_{1}, E_{2}\right\}$ is a split of $G$, and $e \notin E(G)$, then we use $G\left(E_{1} ; e\right), G\left(E_{2} ; e\right)$ to denote the members of the (unique) simple decomposition of $G$ associated with $\left\{E_{1}, E_{2}\right\}$ and $e$.

Lemma 1. If $G$ is a non-separable graph and $G \rightarrow\left\{G_{1}, G_{2}\right\}$, then $G_{1}$ and $G_{2}$ are non-separable.

Proof. Let $\left\{E_{1}, E_{2}\right\}$ be the split and $e$ be the marker associated with $\left\{G_{1}, G_{2}\right\}$. It is easy to see that, for any $A \subseteq E_{1}, W\left(G_{1} ; A\right)=W(G ; A)$. The result follows.

Lemma 2. Let $G \rightarrow\left\{G_{1}, G_{2}\right\}$ with marker $e$, and let $A \subset E_{1}$. Then $\{A, E(G) \backslash A\}$ is a split of $G$ if and only if $\left\{A,\left(E_{1} \backslash A\right) \cup\{e\}\right\}$ is a split of $G_{1}$.

Proof. Again, the result follows from the fact that $W(G ; A)=$ $W\left(G_{1} ; A\right)$.

Lemma 3. Let $\left\{E_{1}, E_{2}\right\}$ and $\left\{E_{3}, E_{4}\right\}$ be splits of $G$ such that $E_{3} \subset E_{1}$.

Let e, $f$ be distinct elements which are not edges of $G$. Then

$$
\begin{aligned}
& G\left(E_{1} ; e\right)\left(E_{3} ; f\right)=G\left(E_{3} ; f\right), \text { and } \\
& G\left(E_{1} ; e\right)\left(\left(E_{1} \backslash E_{3}\right) \cup\{e\} ; f\right)=G\left(E_{4} ; f\right)\left(\left(E_{4} \backslash E_{2}\right) \cup\{f\} ; e\right) .
\end{aligned}
$$

Proof. As before, for any $A \subseteq E_{1}$, we have

$$
\begin{aligned}
W\left(G\left(E_{1} ; e\right) ; A\right)=W(G ; A) \text { and } W\left(G\left(E_{1} ; e\right) ; A\right. & \cup\{e\}) \\
& =W\left(G ; A \cup E_{2}\right)
\end{aligned}
$$

The result follows from these formulae, because $W(H ;\{h\})$ is the set of ends of $h$ in $H$, for any edge $h$ of graph $H$.

Lemma 4. Let $\left\{E_{1}, E_{2}\right\}$ and $\left\{E_{3}, E_{4}\right\}$ be splits of the graph $G$, such that $\left|E_{1} \cap E_{3}\right| \geqq 2$ and $E_{1} \cup E_{3} \neq E(G)$. Then $\left\{E_{1} \cap E_{3}, E_{2} \cup E_{4}\right\}$ is a split of $G$.

Proof. For this proof, we abbreviate $W(G ; A)$ to $W(A)$, for any $A \subseteq E(G)$. Clearly, $W\left(E_{1} \cap E_{3}\right), W\left(E_{2} \cap E_{4}\right) \subseteq W\left(E_{1}\right) \cup W\left(E_{3}\right)$. Any element $v$ of $W\left(E_{1} \cap E_{3}\right) \cap W\left(E_{2} \cap E_{4}\right)$ must be in $W\left(E_{1}\right) \cap$ $W\left(E_{3}\right)$. (If not, suppose that $v \in W\left(E_{1}\right) \backslash W\left(E_{3}\right)$; then $v$ is incident with no member of $E_{4}$, so $v \notin W\left(E_{2} \cap E_{4}\right)$.) Thus

$$
\begin{aligned}
\left|W\left(E_{1} \cap E_{3}\right)\right|+ & \left|W\left(E_{2} \cap E_{4}\right)\right| \leqq\left|W\left(E_{1}\right) \cap W\left(E_{3}\right)\right| \\
& +\left|W\left(E_{1}\right) \cup W\left(E_{3}\right)\right|=\left|W\left(E_{1}\right)\right|+\left|W\left(E_{3}\right)\right|=4 .
\end{aligned}
$$

But we are given that $E_{2} \cap E_{4} \neq \emptyset$, and so, since $G$ is non-separable,

$$
\left|W\left(E_{2} \cap E_{4}\right)\right| \geqq 2 .
$$

The result follows.
We introduce here some further notational conventions. Where $A$ is a set and $e$ is an element, we abbreviate $A \cup\{e\}$ to $A+e$ and $A \backslash\{e\}$ to $A-e$. In the absence of parentheses, set operations are to be performed from the left; for example, $A \backslash B+e$ denotes $(A \backslash B) \cup\{e\}$. For sets $A$ and $B, A$ meets $B$ means that $A \cap B \neq \emptyset$. Partitions $\left\{A_{1}, A_{2}\right\}$ and $\left\{B_{1}, B_{2}\right\}$ of a set $E$ are said to cross if $A_{1} \cap B_{1}, A_{1} \cap B_{2}, A_{2} \cap B_{1}$, and $A_{2} \cap B_{2}$ are all non-empty.

The organization of the paper is as follows. In Sections 2 and 3 we introduce "decomposition frames" and prove the main results concerning them. In Section 4 these results are applied to the decomposition of non-separable graphs, which has already been introduced. In Section 5 we describe a much deeper application, a decomposition theory for families of sets. Section 6 is devoted to a special case of the set-family theory, a decomposition theory for matroids; it is proved that, for matroids, "prime" is equivalent to " 3 -connected". In Section 7 we derive a "substitution" decomposition theory for set families from the theory of Section 5 .

For the special case in which the set families are "clutters" this topic has been previously investigated in the contexts of Boolean functions and simple games.

The theory presented in this paper appeared in the thesis [10] of the first author, which was supervised by the second author. It was inspired by the work of Louis Billera and has benefited, at an earlier stage, from conversations with Professors Billera, J. A. Bondy and R. E. Bixby. The authors are grateful to an anonymous referee for a careful reading of the manuscript which resulted in several improvements, including a significant simplification of the proof of Theorem 15 . This research has been supported by fellowships and grants from the National Research Council of Canada.
2. Decomposition frames: basic properties. In this section we introduce the notion of decomposition frame, and explain its relevance to the graph decomposition material of Section 1. We derive some elementary properties, and prove a rudimentary unique decomposition theorem which is the basis for the deeper results of Section 3.

In absorbing the following definitions, the reader should keep in mind the graph decomposition example of Section 1 . Let $\mathscr{N}$ be a class and $E$ be a function defined on $\mathscr{N}$ such that, for each $N \in \mathscr{N}, E(N)$ is a finite set, called the set of cells of $N$. Let $\rightarrow$ be a relation associating elements $N$ of $\mathscr{N}$ to two-element subsets $\left\{N_{1}, N_{2}\right\}$ of $\mathscr{N}$, written $N \rightarrow\left\{N_{1}, N_{2}\right\}$. The triple $(\mathscr{N}, E, \rightarrow)$ is a decomposition frame if F1, F2, F3, F4 below are satisfied.

F1. If $N \rightarrow\left\{N_{1}, N_{2}\right\}$, then for some $e \notin E(N)$ and some partition $\left\{E_{1}, E_{2}\right\}$ of $E(N)$ with $\left|E_{1}\right| \geqq 2 \leqq\left|E_{2}\right|$, we have $E\left(N_{1}\right)=E_{1}+e$, $E\left(N_{2}\right)=E_{2}+e$.

In the context of $\mathrm{F} 1,\left\{N_{1}, N_{2}\right\}$ is called a simple decomposition of $N$, $e$ is called the marker of the simple decomposition, and $\left\{E_{1}, E_{2}\right\}$ is called the split of $N$ corresponding to the simple decomposition.

F2. For a split $\left\{E_{1}, E_{2}\right\}$ of $N \in \mathscr{N}$ and $e \notin E(N)$, there is exactly one simple decomposition $\left\{N_{1}, N_{2}\right\}$ of $N$ with marker $e$ corresponding to $\left\{E_{1}, E_{2}\right\}$.

Given a split $\left\{E_{1}, E_{2}\right\}$ of $N$ and $e \notin E(N)$, we denote by $N\left(E_{i} ; e\right), i=1$ and 2 , the unique element of $\mathscr{N}$ such that $E\left(N\left(E_{i} ; e\right)\right)=E_{i}+e$ and $N \rightarrow\left\{N\left(E_{1} ; e\right), N\left(E_{2} ; e\right)\right\}$.

F3. Let $\left\{E_{1}, E_{2}\right\}$ be a split of $N \in \mathcal{N}$, let $A \subset E_{1}$, and $e \notin E(N)$. Then $\{A, E(N) \backslash A\}$ is a split of $N$ if and only if $\left\{A,\left(E_{1}+e\right) \backslash A\right\}$ is a split of $N\left(E_{1} ; e\right)$.

F4. Let $\left\{E_{1}, E_{2}\right\},\left\{E_{3}, E_{4}\right\}$ be splits of $N \in \mathscr{N}$ such that $E_{3} \subset E_{1}$, and let $e, f \notin E(N), e \neq f$. Then

$$
\begin{aligned}
& N\left(E_{1} ; e\right)\left(E_{3} ; f\right)=N\left(E_{3} ; f\right), \text { and } \\
& N\left(E_{1} ; e\right)\left(E_{1} \backslash E_{3}+e ; f\right)=N\left(E_{4} ; f\right)\left(E_{4} \backslash E_{2}+f ; e\right) .
\end{aligned}
$$

It follows immediately from the definitions and Lemmas 1, 2, 3 of Section 1 that, where $\mathscr{N}$ is the class of finite non-separable graphs, $E(G)$ denotes the edge-set of $G$, and $\rightarrow$ is as defined in Section 1, $(\mathscr{N}, E, \rightarrow)$ is a decomposition frame. In general, we will refer to the elements of $\mathscr{N}$ as the objects of the decomposition frame. We define the terms decomposition, (strict) refinement, markers (of a decomposition), equivalent, minimal, trivial, and prime for decomposition frames just as they were defined in Section 1. The only differences are the substitution of "cell" for "edge", and "object" for "non-separable graph". The notion of "decomposition tree" also extends to the present context.

A set of splits of an element $N$ of $\mathscr{N}$ is compatible if no two members of the set cross. By (F3) and the definition of decomposition, every decomposition $D$ of $N$ gives rise to a set of compatible splits of $N$, one for each simple refinement in a derivation of $N$. Given a decomposition $D_{i-1}$ of $N$, a split $\left\{A_{i}{ }^{\prime}, B_{i}{ }^{\prime}\right\}$ of a member $N^{\prime}$ of $D$ generates a unique simple refinement $D_{i}$ of $D_{i-1}$, and any other split of $N^{\prime}$ which does not cross $\left\{A_{i}{ }^{\prime}, B_{i}{ }^{\prime}\right\}$ induces, by (F3), a unique split of a member of $D_{i}$. Thus an ordered set $\left\{\left\{A_{i}, B_{i}\right\}: 1 \leqq i \leqq k\right\}$ of compatible splits of $N$ generates a unique decomposition of $N$. In fact, as the next result shows, this decomposition does not depend on the order.

Theorem 2. For $N \in \mathscr{N}$, any set of compatible splits of $N$ generates a unique decomposition of $N$.

Proof. Let $\left\{\left\{A_{i}, B_{i}\right\}: i \in I\right\}$ be a set of compatible splits of $N$. We know that for any fixed ordering of $I$, the decomposition generated is unique. It is true that any ordering of $I$ can be obtained from any other one by a sequence of interchanges of adjacent elements. Thus it suffices to prove that the decomposition $D$ obtained from the ordering $S: i_{1}, i_{2}, \ldots, i_{j-1}, i_{j}, i_{j+1}, i_{j+2}, \ldots, i_{k}$ of $I$ is the same as the decomposition $D^{\prime}$ obtained from the ordering $S^{\prime}: i_{1}, i_{2}, \ldots, i_{j-1}, i_{j+1}, i_{j}, i_{j+2}, \ldots, i_{k}$. Let $D_{0}=D_{0}{ }^{\prime}=\{N\}$ and let $D_{m},\left(D_{m}{ }^{\prime}\right), 1 \leqq m \leqq k$, denote the decomposition of $N$ generated by the splits $\left\{A_{i}, B_{i}\right\}$, where $i$ runs through the first $m$ terms of $S\left(S^{\prime}\right)$ in that order. Then $D_{k}=D$ and $D_{k}{ }^{\prime}=D^{\prime}$. Clearly, $D_{j-1}=D_{j-1}{ }^{\prime}$. If $\left\{A_{j}, B_{j}\right\}$ and $\left\{A_{j+1}, B_{j+1}\right\}$ induce splits in different members of $D_{j-1}$, then clearly $D_{j+1}=D_{j+1}{ }^{\prime}$. If they induce splits of the same member of $D_{j-1}$, then $D_{j+1}=D_{j+1}{ }^{\prime}$, by (F4). In either case $D_{j+1}{ }^{\prime}=D_{j+1}$, and so $D=D^{\prime}$, as required.

In view of the notion of "decomposition tree", with any set of compatible 2 -element partitions of a finite set, Theorem 2 associates a tree structure. Essentially the same concept is an important idea of [13], although the two papers are otherwise quite different.

A split of a member $N$ of $\mathscr{N}$ is said to be good if it is crossed by no other split of $N$. Good splits play an extremely important role in this uniqueness theory. Obviously a prime member of $\mathscr{N}$ can have no good split, but there can exist non-primes having no good split. For example, it is easy to see that, with respect to the graph decomposition described in Section 1, bonds and polygons have no good split.

Theorem 3. Let $N \in \mathscr{N}$. Then $N$ has a unique minimal decomposition, each of whose members has no good split.

Lemma 5. If $\left\{E_{1}, E_{2}\right\}$ is a split of $N$ and $\left\{E_{3}, E_{4}\right\}$ is a good split of $N$ with $E_{3} \subset E_{1}$, then $\left\{E_{3}, E_{1} \backslash E_{3}+e\right\}$ is a good split of $N\left(E_{1} ; e\right)$.

Proof. If the lemma is not true, then there exists a split $\left\{A, E_{1} \backslash A+e\right\}$ of $N\left(E_{1} ; e\right)$ which crosses $\left\{E_{3}, E_{1} \backslash E_{3}+e\right\}$. Then by (F3), $\{A, E(N) \backslash A\}$ is a split of $N$, and it is easy to check that it crosses $\left\{E_{3}, E_{4}\right\}$, contradicting the goodness of $\left\{E_{3}, E_{4}\right\}$.

Lemma 6. If $\left\{E_{1}, E_{2}\right\}$ is a good split of $N \in \mathscr{N}$ and $\left\{E_{3}, E_{1} \backslash E_{3}+e\right\}$ is a good split of $N\left(E_{1} ; e\right)$, then $\left\{E_{3}, E(N) \backslash E_{3}\right\}$ is a good split of $N$.

Proof. By (F3), $\left\{E_{3}, E(N) \backslash E_{3}\right\}$ is a split of $N$. If it is not a good split, then there exists a split $\left\{E_{5}, E_{6}\right\}$ of $N$ which crosses it. Now $\left\{E_{5}, E_{6}\right\}$ cannot cross $\left\{E_{1}, E_{2}\right\}$, since $\left\{E_{1}, E_{2}\right\}$ is good, and cannot be equal to $\left\{E_{1}, E_{2}\right\}$, since $\left\{E_{1}, E_{2}\right\}$ does not cross $\left\{E_{3}, E(N) \backslash E_{3}\right\}$. By interchanging $E_{5}$ with $E_{6}$ if necessary, we may conclude that $E_{5} \subset E_{1}$ or $E_{5} \supseteq E_{1}$. But in the second case, $E_{6} \cap E_{3}=\emptyset$, which would imply that $\left\{E_{5}, E_{6}\right\}$ does not cross $\left\{E_{3}, E(N) \backslash E_{3}\right\}$. Therefore, $E_{5} \subset E_{1}$, so by (F3), $\left\{E_{5}, E_{1} \backslash\right.$ $\left.E_{5}+e\right\}$ is a split of $N\left(E_{1} ; e\right)$. But it follows from the assumption that $\left\{E_{5}, E_{6}\right\}$ crosses $\left\{E_{3}, E(N) \backslash E_{3}\right\}$, that $\left\{E_{5}, E_{1} \backslash E_{5}+e\right\}$ crosses $\left\{E_{3}, E_{1} \backslash\right.$ $\left.E_{3}+e\right\}$, contradicting the goodness of $\left\{E_{3}, E_{1} \backslash E_{3}+e\right\}$.

Proof of Theorem 3. Let $D$ be a decomposition of $N$, and suppose that $D$ is generated by the set $S$ of splits of $N$. If there is a good split of $N$ which is not in $S$ then, by Lemma 5 , there is a member of $D$ having a good split. Therefore, if $D$ has the property that none of its members has a good split, then every good split of $N$ is in $S$. It follows that every such decomposition $D$ is a refinement of the (unique, by Theorem 2) decomposition $D^{\prime}$ of $N$ generated by the set of good splits of $N$. Thus, to prove the theorem, it suffices to show that no member of $D^{\prime}$ has a good split. But this follows from Lemma 6.

At this point we introduce another example of a decomposition frame. Define a split system $N$ to be a finite set $E(N)$ together with a set of
partitions (called splits) $\left\{E_{1}, E_{2}\right\}$ of $E(N)$, such that $\left|E_{1}\right| \geqq 2 \leqq\left|E_{2}\right|$. Given a split $\left\{E_{1}, E_{2}\right\}$ of $N$ and $e \notin E(N)$, define a split system $N\left(E_{i} ; e\right)$, for $i=1$ and 2 , having $E\left(N\left(E_{i} ; e\right)\right)=E_{i}+e$ and having as splits precisely the partitions of the form $\left\{E_{3}, E_{i} \backslash E_{3}+e\right\}$ where $\left\{E_{3}, E(N) \backslash E_{3}\right\}$ is a split of $N$, and $E_{3} \subset E_{i}$. Then we write

$$
N \rightarrow\left\{N\left(E_{1} ; e\right), N\left(E_{2} ; e\right)\right\}
$$

It is easily proved that, where $\mathscr{N}$ is the class of all split systems, $(\mathscr{N}, E, \rightarrow)$ is a decomposition frame. It is, in a sense, the simplest of all decomposition frames, since its members have no structure other than that necessary to satisfy the frame axioms. (This frame does not satisfy the additional properties which are needed for the results of the next section.)

The split system decomposition frame illustrates an interesting aspect of the present theory. Many decomposition theories, such as the theory of prime factorization of integers, are based on possibilities for reversing some uniquely determined composition. Though several instances described here do have this property, others do not, and the split system frame is an obvious example. If $\left\{N_{1}, N_{2}\right\}$ is a simple decomposition of the split system $N$ associated with the split $\left\{E_{1}, E_{2}\right\}$ of $N$, then $\left\{N_{1}, N_{2}\right\}$ determines the splits of $N$ not crossing $\left\{E_{1}, E_{2}\right\}$, but does not generally determine the ones that do cross $\left\{E_{1}, E_{2}\right\}$. Moreover, while the graph decomposition frame of Section 1 does have a unique composition, a slight variant of it does not. Suppose that we define the frame to have as objects equivalence classes of non-separable graphs, where $G$ and $G^{\prime}$ are equivalent if there is an isomorphism from $G$ to $G^{\prime}$ which is the identity on $E(G)$. (This is just a precise way of saying that we want to "forget" vertex names.) Now if we have a simple decomposition $\left\{G_{1}, G_{2}\right\}$ of $G$ with marker $e,\left\{G_{1}, G_{2}\right\}$ is also a simple decomposition of an object $G^{\prime}$ obtained by "identifying the ends of $e$ in the opposite order". Therefore, this new frame does not have a unique composition, but Lemmas 1 to 4 are still true and, as we shall see, these are the results needed to derive Theorem 1. The point that we are making is that a successful decomposition theory can exist in the absence of a uniquely determined composition.
3. Decomposition frames: main theorems. We say that a decomposition frame $(\mathscr{N}, E, \rightarrow)$ has the intersection property if, whenever $\left\{E_{1}, E_{2}\right\}$ and $\left\{E_{3}, E_{4}\right\}$ are splits of $N \in \mathscr{N}$ such that $\left|E_{1} \cap E_{3}\right| \geqq 2$ and $E_{1} \cup E_{3} \neq E(N)$, then $\left\{E_{1} \cap E_{3}, E_{2} \cup E_{4}\right\}$ is a split of $N$. By Lemma 4, the graph decomposition frame of Section 1 has the intersection property. We say that $N \in \mathscr{N}$ is brittle if every partition $\left\{E_{1}, E_{2}\right\}$ of $N$ such that $\left|E_{1}\right|,\left|E_{2}\right| \geqq 2$, is a split of $N$. We say that $N \in \mathscr{N}$ is semi-brittle if $E(N)=\left\{e_{0}, e_{1}, \ldots, e_{n-1}\right\}$, and the splits of $N$ are precisely the partitions of $E(N)$ of the form $\left\{\left\{e_{i+1}, e_{i+2}, \ldots, e_{i+j}\right\},\left\{e_{i+j+1}, e_{i+j+2}, \ldots, e_{i}\right\}\right\}$,
where subscripts are modulo $n$ and $j, n-j \geqq 2$. We observe that, if $N$ is brittle or semi-brittle, then $N$ has no good split. We also observe that, with respect to the graph decomposition frame, bonds are brittle and polygons are semi-brittle.

Theorem 4. Let $(\mathscr{N}, E, \rightarrow)$ be a decomposition frame having the intersection property, and let $N \in \mathcal{N}$. Then $N$ has a unique minimal decomposition, each of whose members is prime, brittle, or semi-brittle.

We say that $(\mathscr{N}, E, \rightarrow)$ has the transitivity property if, whenever

$$
\left\{\left\{e_{1}, e_{2}\right\}, E(N) \backslash\left\{e_{1}, e_{2}\right\}\right\} \text { and }\left\{\left\{e_{2}, e_{3}\right\}, E(N) \backslash\left\{e_{2}, e_{3}\right\}\right\}
$$

are splits of $N \in \mathscr{N}$, then so also is

$$
\left\{\left\{e_{1}, e_{3}\right\}, E(N) \backslash\left\{e_{1}, e_{3}\right\}\right\}
$$

The graph decomposition frame does not have the transitivity property. In fact, it is clear that, if a frame has the transitivity property, then every semi-brittle object $N$ has $E(N) \leqq 3$, and therefore is prime. Therefore, the following important result is an immediate consequence of Theorem 4.

Theorem 5. Let $(\mathscr{N}, E, \rightarrow)$ be a decomposition frame having the intersection and transitivity properties, and let $N \in \mathscr{N}$. Then $N$ has a unique minimal decomposition, each of whose members is prime or brittle.

On the other hand, Theorem 4 follows from Theorem 3, together with the following characterization of objects having no good split.

Theorem 6. Let $(\mathscr{N}, E, \rightarrow)$ be a decomposition frame having the intersection property. An object $N \in \mathscr{N}$ has no good split if and only if $N$ is prime, brittle, or semi-brittle.

Lemma 7. Let $N$ be an object of a decomposition frame having the intersection property, and suppose that $N$ is not prime and has no good split. Then there exists an ordering $e_{0}, e_{1}, \ldots, e_{n-1}$ of $E(N)$ such that, for $0 \leqq i \leqq n-1$,

$$
\left\{\left\{e_{i}, e_{i+1}\right\}, E(N) \backslash\left\{e_{i}, e_{i+1}\right\}\right\} \text { is a split of } N .
$$

(Subscripts are modulo $n$ ).
Proof. The truth of the lemma is easily checked for $|E(N)|=4$ (and for $|E(N)|<4)$. Suppose that it is true for all $N$ such that $|E(N)| \leqq$ $m \geqq 4$, and suppose that we are given $N$ with $|E(N)|=m+1$, not prime and having no good split. Then $N$ has a split $\left\{E_{1}, E_{2}\right\}$. Let $N_{i}$ denote $N\left(E_{i} ; e\right)$ for $i=1$ and 2 , where $e \notin E(N)$. We claim that $N_{1}$ and $N_{2}$ have no good split. Let $\left\{A, E_{1} \backslash A+e\right\}$ be a split of $N_{1}$. Then $\{A, E(N) \backslash A\}$ is a split of $N$ by (F3). There exists a split $\{B, E(N) \backslash B\}$ of $N$ crossing $\{A, E(N) \backslash A\}$. Rename, if necessary, so that $E(N) \backslash B$ meets $E_{1} \backslash A$.

Case 1. $B \subseteq E_{1}$. Then $\left\{B, E_{1} \backslash B+e\right\}$ is a split of $N_{1}$, by (F3). Now $B \cap A \neq \emptyset,\left(E_{1} \backslash A+e\right) \cap\left(E_{1} \backslash B+e\right) \neq \emptyset$, and $A \cap\left(E_{1} \backslash B+e\right) \neq \emptyset$. If $B \cap\left(E_{1} \backslash A+e\right)=\emptyset$, then $B \subseteq A$, a contradiction. Thus $\left\{A, E_{1} \backslash\right.$ $A+e\}$ is not a good split of $N_{1}$.

Case $2 . B \cap E_{2} \neq \emptyset$. Then, by the intersection property, $\left\{E_{1} \backslash B\right.$, $\left.B \cup E_{2}\right\}$ is a split of $N$. Therefore, $\left\{E_{1} \backslash B, E_{1} \cap B+e\right\}$ is a split of $N_{1}$, by (F3). Now $A$ meets $E_{1} \backslash B$ and $E_{1} \cap B+e$, since $\{A, E(N) \backslash A\}$ crosses $\{B, E(N) \backslash B\} ; E_{1} \backslash B$ meets $E_{1} \backslash A+e$, since $E(N) \backslash B$ meets $E_{1} \backslash A ; E_{1} \cap B+e$ meets $E_{1} \backslash A+e$, because $e$ is an element of each. Thus $\left\{A, E_{1} \backslash A+e\right\}$ is not a good split of $N_{1}$.

Therefore, $N_{1}$ has no good split; similarly, $N_{2}$ has no good split. There exists a split of $N$ crossing $\left\{E_{1}, E_{2}\right\}$. It follows from the intersection property and (F3) that, for $i=1$ and $2, N_{i}$ is not prime unless $\left|E_{i}\right|=2$. From this fact, the induction hypothesis, and (F3), we conclude that there exist orderings $c_{1}, c_{2}, \ldots, c_{p}$ of $E_{1}$ and $d_{1}, d_{2}, \ldots, d_{q}$ of $E_{2}$ such that

$$
\left\{\left\{c_{i}, c_{i+1}\right\}, E(N) \backslash\left\{c_{i}, c_{i+1}\right\}\right\} \text { is a split of } N \text { for } i=1,2 \ldots, p-1,
$$

and

$$
\left\{\left\{d_{i}, d_{i+1}\right\}, E(N) \backslash\left\{d_{i}, d_{i+1}\right\}\right\} \text { is a split of } N \text { for } i=1,2, \ldots, q-1 .
$$

It follows from the intersection property that

$$
\left\{\left\{c_{i}, c_{i+1}, \ldots, c_{j}\right\}, E(N) \backslash\left\{c_{i}, c_{i+1}, \ldots, c_{j}\right\}\right\} \text { is a split of } N
$$

$$
\text { whenever } 1 \leqq i<j \leqq p \text {, }
$$

and similarly for the $d_{i}$.
We can choose a split $\left\{E_{3}, E_{4}\right\}$ of $N$ and the orderings $c_{1}, \ldots, c_{p}$ and $d_{1}, \ldots, d_{q}$ so that $\left\{E_{3}, E_{4}\right\}$ crosses $\left\{E_{1}, E_{2}\right\}$ and $\left\{c_{p}, d_{1}\right\} \subseteq E_{3}$. (Proof. If not, there exists such a split $\left\{E_{3}{ }^{\prime}, E_{4}{ }^{\prime}\right\}$ with $c_{1}, c_{p} \in E_{3}{ }^{\prime}$ and $d_{1}, d_{q} \in E_{4}{ }^{\prime}$. Then $\left|E_{1}\right|>2$, and so by the intersection property $\left\{E_{1}-c_{p}, E_{2}+c_{p}\right\}$ is a split of $N$. Then

$$
\left\{E_{3}, E_{4}\right\}=\left\{E_{3}^{\prime} \cap\left(E_{2}+c_{p}\right), E_{4}^{\prime} \cup\left(E_{1}-c_{p}\right)\right\}
$$

is a split having the required property.) Now, by the intersection property, the partitions $\left\{E_{5}, E_{6}\right\}$ and $\left\{E_{7}, E_{8}\right\}$ of $E(N)$, given by

$$
E_{5}=E_{3} \cap\left(E_{1}+d_{1}\right) \text { and } E_{7}=E_{5} \cap\left(E_{2}+c_{p}\right),
$$

are splits of $N$. That is, $\left\{\left\{c_{p}, d_{1}\right\}, E(N) \backslash\left\{c_{p}, d_{1}\right\}\right.$ is a split of $N$. It is easy to use this fact and the intersection property to show that $\left\{\left\{c_{1}, d_{q}\right\}\right.$, $\left.E(N) \backslash\left\{c_{1}, d_{q}\right\}\right\}$ is also a split of $N$. Therefore, $c_{1}, c_{2}, \ldots, c_{p}, d_{1}, \ldots, d_{q}$ is the required ordering of $E(N)$, and the lemma is proved by induction.

Proof of Theorem 6. Suppose that $N$ has no good split and is neither prime nor semi-brittle. Then there is an ordering $e_{0}, e_{1}, \ldots, e_{n-1}$ of $E(N)$ as in Lemma 7, and there exist a split $\left\{E_{1}, E_{2}\right\}$ of $N$ and integers $i, j, k$ with $0<i<j<k<n$ such that $e_{0}, e_{j} \in E_{1}$ and $e_{i}, e_{k} \in E_{2}$.

Applying the intersection property, we obtain a split $\left\{E_{3}, E_{4}\right\}$ where

$$
E_{3}=E_{1} \cap\left\{e_{0}, e_{1}, \ldots, e_{j}\right\} ;
$$

applying it again, we obtain a split $\left\{E_{5}, E_{6}\right\}$ where

$$
E_{5}=E_{3} \cap\left\{e_{j}, e_{j+1}, \ldots, e_{n-1}, e_{0}\right\}=\left\{e_{0}, e_{j}\right\} .
$$

Now let $m$ be an integer such that $2 \leqq m \leqq j-1$. By the intersection property, $\left\{E_{7}, E_{8}\right\}$ is a split of $N$, where

$$
E_{8}=E_{2} \cup\left\{e_{1}, e_{2}, \ldots, e_{j-1}\right\} .
$$

Again, $\left\{E_{9}, E_{10}\right\}$ is a split, where

$$
E_{9}=E_{7} \cup\left\{e_{m}, e_{m+1}, \ldots, e_{j}\right\}
$$

If we repeat the argument of the first part of the proof, with $\left\{E_{9}, E_{10}\right\}$ replacing $\left\{E_{1}, E_{2}\right\}, e_{1}$ replacing $e_{i}$, and $e_{m}$ replacing $e_{j}$, we can conclude that $\left\{\left\{e_{0}, e_{m}\right\}, E(N) \backslash\left\{e_{0}, e_{m}\right\}\right\}$ is a split of $N$. A similar argument can be applied to the case where $j+1 \leqq m \leqq n-2$. It follows that, for $1 \leqq m \leqq n-1, \quad\left\{\left\{e_{0}, e_{m}\right\}, E(N) \backslash\left\{e_{0}, e_{m}\right\}\right\}$ is a split of $N$. Now let $\left\{E_{1}, E_{2}\right\}$ be any partition of $N$ such that $\left|E_{1}\right| \geqq 2 \leqq\left|E_{2}\right|$; we may assume that $e_{0} \in E_{1}$. Using the splits $\left\{\left\{e_{0}, e_{j}\right\}, E(N) \backslash\left\{e_{0}, e_{j}\right\}\right\}$ for $e_{j} \in E_{1}-e_{0}$ and applying the intersection property repeatedly, we conclude that $\left\{E_{1}, E_{2}\right\}$ is a split of $N$. Therefore, $N$ is brittle, and the proof is complete.
4. Graph decomposition. In this section we apply the theory of the last two sections to prove Theorem 1. We also discuss some other aspects of the graph decomposition theory.

Proof of Theorem 1. It is a consequence of Lemmas 1 to 4 that, where $\mathscr{G}$ is the class of non-separable graphs, $E$ means "edge-set", and $\rightarrow$ is as defined in Section 1, ( $\mathscr{G}, E, \rightarrow$ ) is a decomposition frame having the intersection property. Now let $G \in \mathscr{G}$ and let $e, f \in E(G)$, where $|E(G)| \geqq 4$. Then $\{\{e, f\}, E(G) \backslash\{e, f\}\}$ is a split of $G$ if and only if $e$ and $f$ are in parallel (have the same ends in $G$ ) or $e$ and $f$ are in series (are the two edges incident with a vertex of degree 2 in $G$ ). Moreover, if $e$ and $f$ are in parallel in $G$, and $f$ and $g$ are in parallel in $G$, then $e$ and $g$ are in parallel in $G$. On the other hand, if $e$ and $f$ are in series in $G$, and $f$ and $g$ are in series in $G$, then $e$ and $g$ are not in series in $G$. Finally, it is not possible for both $e$ and $f$ to be in parallel and $f$ and $g$ to be in series. From these observations we deduce that if $G$ is brittle, every two edges of $G$ are in parallel, so that $G$ is a bond. Also, if $G$ is semi-brittle, then there is an ordering $e_{0}, e_{1}, \ldots, e_{n-1}$ of $E(G)$ such $e_{i}$ is in series with $e_{i+1}$ for $0 \leqq i \leqq$ $n-1$, so that $G$ is a polygon. Therefore, Theorem 1 follows immediately from Theorem 4.

Given a non-separable graph $G$, let us call the decomposition of $G$ whose uniqueness is asserted in Theorem 1 the standard decomposition of $G$. A consequence of Theorem 1 is that every decomposition of $G$ consisting of 3 -connected graphs, polygons, and bonds is a refinement of the standard decomposition $D$ of $G$. Any such decomposition $D^{\prime}$ must result from replacing each bond and polygon of $D$ by the members of a decomposition of the bond or polygon. If $D^{\prime} \neq D$, this means that $D^{\prime}$ has two polygons or two bonds which share a marker. On the other hand, it is easy to show, using the fact that $D$ is generated by the good splits of $G$ (in any order), that $D$ cannot have two such bonds or polygons. Therefore, we can avoid the "minimal" in Theorem 1 by stating it thus: Every non-separable graph has a unique decomposition consisting of 3 -connected graphs, polygons, and bonds with the property that no two bonds and no two polygons share a marker.

According to a notion of 3 -connectivity which is perhaps more common than Tutte's, a non-separable graph $G$ is 3 -connected if and only if it has no "Whitney split"; that is, there does not exist a partition $\left\{E_{1}, E_{2}\right\}$ of $E(G)$ such that

$$
\left|W\left(G ; E_{1}\right)\right|=2 \text { and }\left|V\left(G\left(E_{1}\right)\right)\right|,\left|V\left(G\left(E_{2}\right)\right)\right| \geqq 3 .
$$

Every Whitney split is a Tutte split, and a Tutte split $\left\{E_{1}, E_{2}\right\}$ is a Whitney split if and only if neither $G\left(E_{1}\right)$ nor $G\left(E_{2}\right)$ is a bond. One might want to develop a decomposition theory for non-separable graphs using this more restrictive notion of split, together with the same notion of simple decomposition. However, this cannot be done in the context of decomposition frames, because (F3) is violated; moreover, Whitney splits do not satisfy the intersection property.

In [23, Chapter 11], W. T. Tutte describes a decomposition of a nonseparable graph $G$, whose members are called the "cleavage units" of $G$. Tutte defines this decomposition (uniquely) by restricting simple decompositions to splits satisfying certain additional requirements. He then proves that the resulting decomposition has a number of attractive properties; in particular, its members are 3 -connected graphs, polygons, and bonds, and (though Tutte does not state this explicitly) no two bonds and no two polygons share a marker. It follows from these results and Theorem 1, that the cleavage units of $G$ are precisely the members of the standard decomposition of $G$. Therefore, Tutte's work and our own produce the same canonical decomposition of $G$, but the theorems are different. While Tutte defines the decomposition and establishes some of its properties, we prove that it is characterized by certain of these properties.

Hopcroft and Tarjan [14], [15] have discovered Theorem 1 independently, and have applied it to extend an algorithm for isomorphism of planar 3-connected graphs to an algorithm for isomorphism of arbitrary
planar graphs. (The uniqueness theorem stated in [14] is not correct, but it has been corrected in [15], and a proof appears in an unpublished version of [15]). They also present an algorithm [15] for computing the standard decomposition of a non-separable graph $G$, for which the amount of computation is bounded by a linear function of $|E(G)|$.

In yet another approach to the decomposition of non-separable graphs, MacLane [17] uses the following notion of simple decomposition. Let $G$ be a non-separable graph which is not a polygon and let $\left\{E_{1}, E_{2}\right\}$ be a split of $G$ such that each of $G\left(E_{1}\right), G\left(E_{2}\right)$ contains a circuit. Where $W\left(G ; E_{1}\right)=\{u, v\}$, let $P_{i}$ be a simple path joining $u$ to $v$ in $G\left(E_{3-i}\right)$ for $i=1$ and 2 . Let $G_{i}=G\left(E_{i} \cup E\left(P_{i}\right)\right)$ for $i=1$ and 2 , and call $\left\{G_{1}, G_{2}\right\}$ a simple decomposition of $G$. As usual, we define a (general) decomposition by iterating the simple decomposition. Clearly, a graph will be prime with respect to this notion of decomposition if and only if it is "nodally 3 -connected" [23]. An atom of $G$ is a member of a prime decomposition of $G$ which is not homeomorphic to a bond. MacLane's theorem is that the atoms of $G$ are unique up to a homeomorphism which is the identity on vertices of degree at least three. It is clear that the atoms of $G$ are homeomorphic in this way to the set of non-bond, non-polygon members of our standard decomposition of $G$, and thus that MacLane's theorem is a consequence of Theorem 1.

Finally, we have recently learned, from T. R. S. Walsh, of a theorem of Trakhtenbrot [22] which is closely related to Theorem 1. (The approach taken in [22] is that of "substitution decomposition" as in Section 7.) The paper [25] of Walsh contains a translation of Trakhtenbrot's proof, while in [26], Walsh has used this theorem to derive (independently) Theorem 1.
5. Set-family decomposition. A set family $H$ is a pair $(E, \mathscr{F})$, where $E$ is a finite set of cells of $H$ and $\mathscr{F}$ is a set of non-empty subsets of $E$. A cell $e$ of $H$ such that $\{e\} \in \mathscr{F}$ is called a loop of $H$. A subset $A$ of $E$ is a separator of $H$ if no member of $\mathscr{F}$ meets both $A$ and $E \backslash A$. We say that $H$ is non-separable if its only separators are $E$ and $\emptyset$. If $A \subseteq E$, then the restriction of $H$ to $E \backslash A$ is $H \backslash A=(E \backslash A, \mathscr{F} \backslash A)$ where $\mathscr{F} \backslash A$ denotes $\{F \in \mathscr{F}: F \subseteq E \backslash A\}$. If $e \in E$, we may abbreviate $H \backslash\{e\}$ to $H \backslash e$ and $\mathscr{F} \backslash\{e\}$ to $\mathscr{F} \backslash e$.

Let $\mathscr{H}$ denote the class of non-separable, loopless set families. (We explain these restrictions later. In the terminology of [2], the members of $\mathscr{H}$ are "finite, simple, loopless, connected hypergraphs".) If $H=$ $(E, \mathscr{F}) \in \mathscr{H}$ and $H_{i}=\left(E_{i}+e, \mathscr{F}_{i}\right) \in \mathscr{H}$ for $i=1$ and 2 , where $\left\{E_{1}, E_{2}\right\}$ is a partition of $E$ such that $\left|E_{1}\right| \geqq 2 \leqq\left|E_{2}\right|$ and $e \notin E$, we define $\rightarrow$ by $H \rightarrow\left\{H_{1}, H_{2}\right\}$ if and only if

$$
\begin{array}{r}
\mathscr{F}=\left(\mathscr{F}_{1} \backslash e\right) \cup\left(\mathscr{F}_{2} \backslash e\right) \cup\left\{F_{1} \cup F_{2}-e: e \in F_{1} \in \mathscr{F}_{1},\right. \\
\left.e \in F_{2} \in \mathscr{F}_{2}\right\} ;
\end{array}
$$

in this situation $\left\{E_{1}, E_{2}\right\}$ is said to be a split of $H$. For $H \in \mathscr{H}$, let $E(H)$ denote the set of cells of $H$.

To provide an example of the above simple decomposition, we return to graphs. Given a finite graph $G$, we can associate with $G$ a set family $P M(G)$, called the polygon matroid of $G ; P M(G)=(E(G), \mathscr{F})$, where the members of $\mathscr{F}$ are the edge-sets of simple circuits of $G$. If $G$ has no isolated vertices, then it is a consequence of a result of [28] that $P M(G)$ is a non-separable set family if and only if $G$ is a non-separable graph. Now let $G$ be a non-separable graph having a (graph) split $\left\{E_{1}, E_{2}\right\}$. It is easy to see that $\left\{E_{1}, E_{2}\right\}$ is also a split of $P M(G)$, and that

$$
P M(G) \rightarrow\left\{P M\left(G\left(E_{1} ; e\right)\right), P M\left(G\left(E_{2} ; e\right)\right)\right\}
$$

At this point, it may appear that the set family decomposition directly generalizes the graph decomposition; that this is not the case is demonstrated by two closely-related facts. First, two non-isomorphic nonseparable graphs can have the same polygon matroid; second, the polygon matroid of a non-separable graph $G$ can have (set family) splits which are not (graph) splits of $G$. Nevertheless, there is an extremely close relationship between the set family decomposition theory for $P M(G)$, and the graph decomposition theory for $G$, and we will return to this in the next section.

We now present the main results of the set family decomposition theory.

Theorem 7. $(\mathscr{H}, E, \rightarrow)$ is a decomposition frame, which has the intersection and transitivity properties.

Theorem 8. Each set family $H \in \mathscr{H}$ has a unique minimal decomposition, each of whose members is prime or brittle.

Theorem 8, of course, follows immediately from Theorems 5 and 7. We can also strengthen Theorem 8 by characterizing the brittle set families; they are of a few simple types, which we now describe. Let $H=(E, \mathscr{F})$ be a set family. Then $H$ is a bond if

$$
\mathscr{F}=\{F \subseteq E:|F|=2\} ;
$$

$H$ is a star if

$$
\widetilde{F}=\left\{\left\{e, e^{\prime}\right\}: e \in E-e^{\prime}\right\} \text { for some } e^{\prime} \in E ;
$$

$H$ is a $k$-superstar if

$$
\mathscr{F}=\{F \subseteq E: A \subseteq F,|F| \geqq 2\} \text { for some } A \subseteq E,|A|=k
$$

An $|E|$-superstar is also called a polygon. A superstar is a set family which is a $k$-superstar for some $k$. Theorem 9 below provides a classification of
the brittle set families, and Theorem 10, the main uniqueness theorem for set-family decomposition, is a consequence of Theorems 8 and 9 .

Theorem 9. Set family $H \in \mathscr{H}$ is brittle if and only if $H$ is a bond, a star, or a superstar.

Theorem 10. Each $H \in \mathscr{H}$ has a unique minimal decomposition, each of whose members is a prime, a bond, a star, or a superstar.

An important special case of the present theory, which in fact was developed earlier, occurs when attention is restricted to clutters: set families $(E, \mathscr{F})$ in which no member of $\mathscr{F}$ contains another. The truth of the following proposition is easy to check, and the resulting theorem (Theorem 11) is a consequence of Theorem 10, when we observe which brittle members of $\mathscr{H}$ are clutters.

Proposition 1. If $H \in \mathscr{H}$ and $H \rightarrow\left\{H_{1}, H_{2}\right\}$, then $H$ is a clutter if and only if $H_{1}$ and $H_{2}$ are clutters.

Theorem 11. Each clutter $H \in \mathscr{H}$ has a unique minimal decomposition, each of whose members is a prime, a bond, a star, or a polygon.

If $H_{1}=\left(E_{1}, \mathscr{F}_{1}\right)$ and $H_{2}=\left(E_{2}, \mathscr{F}_{2}\right)$ are set families such that $E_{1} \cap E_{2}=\emptyset$, then the direct sum $H_{1} \oplus H_{2}$ of $H_{1}$ and $H_{2}$ is the set family $\left(E_{1} \cup E_{2}, \mathscr{F}_{1} \cup \mathscr{F}_{2}\right)$. This composition is clearly associative and commutative. The next result is an easy consequence of the definitions.

Proposition 2. The set $A \subseteq E$ is a separator of $H=(E, \mathscr{F})$ if and only if $H=(H \backslash A) \oplus(H \backslash(E \backslash A))$.

It is easy to see that the complement of a separator is a separator, and that the intersection or union of separators is a separator. Thus the minimal non-empty separators (elementary separators) of $H$ partition $E$. The restrictions of $H$ to its elementary separators are called the components of $H$; clearly, they are non-separable. A set family $H=$ $(E, \mathscr{F})$ is said to be null if $E=\emptyset$. The following theorem is easy to derive from the above remarks.

Theorem 12. Each non-null set family $H$ has a unique expression as the direct sum of non-null, non-separable set families.

The existence of the elementary theory of direct sum decomposition justifies the restriction of our theory to non-separable set families. The exclusion of loops is partly explained by observing that, if the marker $e$ were allowed to be a loop in $H_{1}$ or $H_{2}$, the result below would not be true.

Lemma 8. If $\left\{H_{1}, H_{2}\right\}$ is a simple decomposition of $H \in \mathscr{H}$, where $H_{i}=\left(E_{i}+e, \mathscr{F}_{i}\right)$ for $i=1$ and 2 , then $H_{i}$ is uniquely determined from $H, E_{i}$, and $e$ by the formula:

$$
\mathscr{F}_{i}=\left(\mathscr{F} \backslash E_{3-i}\right) \cup\left\{F \cap E_{i}+e: F \in \mathscr{F} \text { meets } E_{1} \text { and } E_{2}\right\} .
$$

Proof. Suppose that $F_{1} \in \mathscr{F}_{1}$. If $e \notin F_{1}$, then $F_{1} \in \mathscr{F}$, so $F_{1} \in \mathscr{F} \backslash E_{2}$. If $e \in F_{1}$, then (since $H_{1}$ is loopless and non-separable) there exists $F_{2} \in \mathscr{F}_{2}$ with $e \in F_{2} \neq\{e\}$. Then $F=F_{1} \cup F_{2}-e \in \mathscr{F}, F$ meets $E_{1}$ and $E_{2}$, and $F_{1}=F \cap E_{1}+e$. Thus

$$
\mathscr{F}_{1} \subseteq\left(\mathscr{F} \backslash E_{2}\right) \cup\left\{F \cap E_{1}+e: F \in \mathscr{F} \text { meets } E_{1} \text { and } E_{2}\right\} .
$$

Now let $F_{1}$ be a member of $\left(\mathscr{F} \backslash E_{2}\right) \cup\left\{F \cap E_{1}+e: F \in \mathscr{F}\right.$ meets $E_{1}$ and $\left.E_{2}\right\}$. If $e \notin F_{1}$, then $F_{1} \in \mathscr{F} \backslash E_{2}$, so $F_{1} \in \mathscr{F}_{1}$ (since $e$ is not a loop of $H_{2}$ ). If $e \in \mathscr{F}_{1}$, we may choose $F \in \mathscr{F}$ meeting $E_{1}$ and $E_{2}$ such that $F_{1}=F \cap E_{1}+e$. By definition of simple decomposition, $F=$ $F_{1}{ }^{\prime} \cup F_{2}{ }^{\prime}-e$, where $e \in F_{i}{ }^{\prime} \in \mathscr{F}_{i}$, for $i=1$ and 2. Then $F_{1}=F_{1}{ }^{\prime}$, so $F_{1} \in \mathscr{F}_{1}$.

The following useful characterization of splits of set families is a consequence of Lemma 8 and the definitions.
Lemma 9. Partition $\left\{E_{1}, E_{2}\right\}$ of $E$ is a split of $H=(E, \mathscr{F}) \in \mathscr{H}$ if and only if $\left|E_{1}\right| \geqq 2 \leqq\left|E_{2}\right|$, and whenever $F_{1}, F_{2} \in \mathscr{F}$ meet both $E_{1}$ and $E_{2}$, then

$$
\left(F_{1} \cap E_{1}\right) \cup\left(F_{2} \cap E_{2}\right) \in \mathscr{F} .
$$

We have explained why loops are not allowed as marker elements. By Lemma 9, the presence of a loop in $H$ will not affect what simple decompositions $H$ can have. Therefore, for consistency, we do not allow loops at all. (We note that, if $H$ is a non-separable clutter having more than one cell, $H$ is necessarily loopless.) We thus arrive at the triple $(\mathscr{H}, E, \rightarrow)$ described above. We begin the proofs of the theorems by verifying that the decomposition frame axioms are satisfied.

Lemma 10. If $H \rightarrow\left\{H_{1}, H_{2}\right\}$, where $H=(E, \mathscr{F}), H_{1}=\left(E_{1}+e, \mathscr{F}_{1}\right)$, and $H_{2}=\left(E_{2}+e, \mathscr{F}_{2}\right)$, and $\left\{E_{3}, E_{4}\right\}$ is a partition of $E$ such that $E_{3} \subset E_{1}$, then $\left\{E_{3}, E_{4}\right\}$ is a split of $H$ if and only if $\left\{E_{3}, E_{1} \backslash E_{3}+e\right\}$ is a split of $H_{1}$.

Proof. We begin by observing that $\left|E_{3}\right| \geqq 2 \leqq\left|E_{4}\right|$ if and only if $\left|E_{3}\right| \geqq 2 \leqq\left|E_{1} \backslash E_{3}+e\right|$. Suppose that $\left\{E_{3}, E_{1} \backslash E_{3}+e\right\}$ is a split of $H_{1}$, and let $F_{1}, F_{2} \in \mathscr{F}$ meet $E_{3}$ and $E_{4}$. Let $F_{3}=\left(F_{1} \cap E_{3}\right) \cup\left(F_{2} \cap E_{4}\right)$. For $i=1$ and 2, let $F_{i}{ }^{\prime}=F_{i}$ if $F_{i} \subseteq E_{1}$ and let $F_{i}{ }^{\prime}=F_{i} \cap E_{1}+e$ otherwise. Then $F_{1}{ }^{\prime}, F_{2}{ }^{\prime} \in \mathscr{F}{ }_{1}$. By assumption,

$$
F_{3}^{\prime}=\left(F_{1}{ }^{\prime} \cap E_{3}\right) \cup\left(F_{2}^{\prime} \cap\left(E_{1} \backslash E_{3}+e\right)\right) \in \mathscr{F}_{1} .
$$

If $e \notin F_{3}{ }^{\prime}$, then $F_{3}=F_{3}{ }^{\prime} \in \mathscr{F}$. Otherwise,

$$
F_{3}=\left(\left(F_{2} \cap E_{2}+e\right) \cup F_{3}{ }^{\prime}\right)-e \in \mathscr{F} .
$$

Thus $\left\{E_{3}, E_{4}\right\}$ is a split of $H$.

Now suppose that $\left\{E_{3}, E_{4}\right\}$ is a split of $H$. For every $F \in \mathscr{F}$ meeting $E_{3}$ and $E_{4}$, there exists $F^{\prime} \in \mathscr{F}_{1}$ meeting both $E_{3}$ and $E_{1} \backslash E_{3}+e$, such that $F \cap E_{3}=F^{\prime} \cap E_{3}$. (Choose $F^{\prime}=F$ if $F \subseteq E_{1}$, and $F^{\prime}=F \cap$ $E_{1}+e$, otherwise.) Moreover, every $F^{\prime} \in \mathscr{F}_{1}$ meeting $E_{3}$ and $E_{1} \backslash E_{3}+e$ arises in this way from some $F$ (there may be several). Let $F_{1}{ }^{\prime}, F_{2}{ }^{\prime}$ be arbitrary members of $\mathscr{F}_{1}$ meeting $E_{3}$ and $E_{1} \backslash E_{3}+e$. Then

$$
\begin{aligned}
& \left(F_{1}^{\prime} \cap E_{3}\right) \cup\left(F_{2}^{\prime} \cap\left(E_{1} \backslash E_{3}+e\right)\right) \\
& =\left(F_{1} \cap E_{3}\right) \cup\left(F_{2}^{\prime} \cap\left(E_{1} \backslash E_{3}+e\right)\right) \\
& \quad=\left(\left(F_{1} \cap E_{3}\right) \cup\left(F_{2} \cap E_{4}\right)\right)^{\prime} \in \mathscr{F}_{1} .
\end{aligned}
$$

Thus $\left\{E_{3}, E_{1} \backslash E_{3}+e\right\}$ is a split of $H_{1}$.
Lemma 11. Let $\left\{E_{1}, E_{2}\right\},\left\{E_{3}, E_{4}\right\}$ be splits of $H \in \mathscr{H}$ with $E_{3} \subset E_{1}$. Then

$$
\begin{aligned}
& H\left(E_{1} ; e\right)\left(E_{3} ; f\right)=H\left(E_{3} ; f\right) \text { and } \\
& H\left(E_{1} ; e\right)\left(E_{1} \backslash E_{3}+e ; f\right)=H\left(E_{4} ; f\right)\left(E_{1} \backslash E_{3}+f ; e\right)
\end{aligned}
$$

Proof. $\mathscr{F}\left(E_{1} ; e\right)$, et cetera, are defined as expected. Now

$$
\begin{aligned}
& \mathscr{F}\left(E_{1} ; e\right)\left(E_{3} ; f\right) \\
& =\left\{F: F \in \mathscr{F}\left(E_{1} ; e\right), F \subseteq E_{3}\right) \cup\left\{F \cap E_{3}+f: F \in \mathscr{F}\left(E_{1} ; e\right)\right. \\
& \left.F \text { meets } E_{1} \text { and } E_{1} \backslash E_{3}+e\right\} \\
& =\left\{F: E_{3} \supseteq F \in \mathscr{F}\right\} \cup\left\{F \cap E_{3}+f: F \in \mathscr{F},\right. \\
& \left.\quad F \text { meets } E_{3} \text { and } E_{1} \backslash E_{3} \text { but not } E_{2}\right\} \cup \\
& \left\{\left(\left(\left(F \cap E_{1}\right)+e\right) \cap E_{3}\right)+f: F \in \mathscr{F}, F \text { meets } E_{3} \text { and } E_{2}\right\} \\
& =\left\{F: E_{3} \supseteq F \in \mathscr{F}\right\} \cup\left\{F \cap E_{3}+f: F \in \mathscr{F}, F \text { meets } E_{3} \text { and } E_{4}\right\} \\
& =\mathscr{F}\left(E_{3} ; f\right) .
\end{aligned}
$$

This proves the first part.
Now

$$
\begin{gathered}
\mathscr{F}\left(E_{1} ; e\right)\left(E_{1} \backslash E_{3}+e ; f\right)=\left\{F: E_{1} \backslash E_{3} \supseteq F \in \mathscr{F}\left(E_{1} ; e\right)\right\} \cup \\
\left\{\left(F \cap\left(E_{1} \backslash E_{3}+e\right)\right)+f: F \in \mathscr{F}\left(E_{1} ; e\right), F \text { meets } E_{3} \text { and } E_{1} \backslash E_{3}+e\right\} \\
=\left\{F: E_{1} \backslash E_{3} \supseteq F \in \mathscr{F}\right\} \cup\left\{\left(F \cap\left(E_{1} \backslash E_{3}\right)\right)+e: F \in \mathscr{F},\right. \\
\left.F \text { meets } E_{1} \backslash E_{3} \text { and } E_{2}, \text { but not } E_{3}\right\} \cup \\
\left\{\left(\left(F \cap\left(E_{1} \backslash E_{3}\right)\right)+f: F \in \mathscr{F}, F \text { meets } E_{1} \backslash E_{3} \text { and } E_{3} \text { but not } E_{2}\right\} \cup\right. \\
\left\{\left(F \cap\left(E_{1} \backslash E_{3}\right)\right)+f+e: F \in \mathscr{F}, F \text { meets } E_{3} \text { and } E_{2}\right\} .
\end{gathered}
$$

By symmetry, $\mathscr{F}\left(E_{4} ; f\right)\left(E_{1} \backslash E_{3}+f ; e\right)$ is obtained from the last expression by interchanging $E_{3}$ with $E_{2}, E_{4}$ with $E_{1}$, and $e$ with $f$. Since this does not change the expression, the result follows.

Combining Lemmas 8, 10, and 11, we obtain the following result.
Theorem 13. $(\mathscr{H}, E, \rightarrow)$ is a decomposition frame.

To complete the proof of Theorem 7, and thus of Theorem 8, we must prove that $(\mathscr{H}, E, \rightarrow)$ has the intersection and transitivity properties. Neither of these results seems to be easy to prove, although the transitivity proof is fairly straightforward.

Theorem 14. $(\mathscr{H}, E, \rightarrow)$ has the intersection property.
Proof. Let $\left\{E_{1}, E_{2}\right\}$ and $\left\{E_{3}, E_{4}\right\}$ be splits of $H=(E, \mathscr{F}) \in \mathscr{H}$, such that $E_{1} \cup E_{3} \neq E$, and $\left|E_{1} \cap E_{3}\right| \geqq 2$. Let $F_{1}, F_{2}$ be members of $\mathscr{F}$ meeting both $E_{1} \cap E_{3}$ and $E_{2} \cup E_{4}$. We must show that

$$
\left(F_{1} \cap E_{1} \cap E_{3}\right) \cup\left(F_{2} \cap\left(E_{2} \cup E_{4}\right)\right) \in \mathscr{F} .
$$

Case 1. $F_{1} \cup F_{2} \nsubseteq E_{1} \cup E_{3}$.
Case 1a. $F_{2} \cap E_{2}, F_{2} \cap E_{4} \neq \emptyset$. Then

$$
\begin{aligned}
&\left(\left(\left(F_{1} \cap E_{1}\right) \cup\left(F_{2} \cap E_{2}\right)\right) \cap E_{2}\right) \cup\left(\left(\left(F_{1} \cap E_{3}\right)\right.\right. \\
&\left.\left.\cup\left(F_{2} \cap E_{4}\right)\right) \cap E_{1}\right) \in \mathscr{F} .
\end{aligned}
$$

But this set is

$$
\begin{aligned}
&\left(F_{2} \cap E_{2}\right) \cup\left(F_{1} \cap E_{1} \cap E_{3}\right) \cup\left(F_{2} \cap E_{1} \cap E_{4}\right)= \\
&\left(F_{1} \cap E_{1} \cap E_{3}\right) \cup\left(F_{2} \cap\left(E_{2} \cup E_{4}\right)\right)
\end{aligned}
$$

as required.
Case 1b. $F_{2} \cap E_{4}=\emptyset$. Then $F_{1} \cap E_{2}, F_{2} \cap E_{2} \neq \emptyset$, since $F_{1}$ and $F_{2}$ meet $E_{2} \cup E_{4}$ and $F_{1} \cup F_{2}$ meets $E_{2} \cap E_{4}$.

Case $1 \mathrm{~b}(\mathrm{i}) . F_{1} \cap E_{1} \cap E_{4}=\emptyset$. Then

$$
\begin{aligned}
\left(F_{1} \cap E_{1} \cap E_{3}\right) & \cup\left(F_{2} \cap\left(E_{2} \cup E_{4}\right)\right) \\
= & \left(F_{1} \cap E_{1} \cap E_{3}\right) \cup\left(F_{2} \cap E_{2}\right) \\
& =\left(F_{1} \cap E_{1}\right) \cup\left(F_{2} \cap E_{2}\right) \in \mathscr{F} .
\end{aligned}
$$

Case 1b(ii). $F_{1} \cap E_{1} \cap E_{4} \neq \emptyset$. Then $F_{3}=\left(F_{1} \cap E_{2}\right) \cup\left(F_{2} \cap E_{1}\right)$ $\in \mathscr{F}$, and $F_{3}$ meets $E_{3}, E_{4}$. Thus $F_{4}=\left(F_{1} \cap E_{3}\right) \cup\left(F_{3} \cap E_{4}\right) \in \mathscr{F}$, and we show that $F_{4}$ meets $E_{2}$ as follows. One of $F_{1}, F_{2}$ meets $E_{2} \cap E_{4}$, but $F_{2} \cap E_{4}=\emptyset$, so $F_{1} \cap E_{2} \cap E_{4} \neq \emptyset$. Therefore, the set $F_{5}=$ $\left(F_{4} \cap E_{1}\right) \cup\left(F_{2} \cap E_{2}\right) \in \mathscr{F}$. But

$$
\begin{aligned}
& F_{5}=\left(F_{1} \cap E_{1} \cap E_{3}\right) \cup\left(F_{3} \cap E_{1} \cap E_{4}\right) \cup\left(F_{2} \cap E_{2}\right) \\
&=\left(F_{1} \cap E_{1} \cap E_{3}\right) \cup\left(F_{2} \cap E_{1} \cap E_{4}\right) \cup\left(F_{2} \cap E_{2}\right) \\
&=\left(F_{1} \cap E_{1} \cap E_{3}\right) \cup\left(F_{2} \cap\left(E_{2} \cup E_{4}\right)\right),
\end{aligned}
$$

as required.
Case 1c. $F_{2} \cap E_{2}=\emptyset$. This case is similar to Case 1b.
Case 2. $F_{1}, F_{2} \subseteq E_{1} \cup E_{3}$.
Case 2a. There exists $F_{3} \in \mathscr{F}$ meeting both $E_{1} \cap E_{3}$ and $E_{2} \cap E_{4}$. Then we define

$$
F_{4}= \begin{cases}\left(F_{1} \cap E_{1}\right) \cup\left(F_{3} \cap E_{2}\right), & \text { if } F_{1} \cap E_{2} \neq \emptyset ; \\ \left(F_{1} \cap E_{3}\right) \cup\left(F_{3} \cap E_{4}\right), & \text { otherwise. }\end{cases}
$$

Then $F_{4} \in \mathscr{F}$ and $F_{4} \cap E_{1} \cap E_{3}=F_{1} \cap E_{1} \cap E_{3}$. Also,

$$
\left(F_{4} \cap E_{1} \cap E_{3}\right) \cup\left(F_{2} \cap\left(E_{2} \cup E_{4}\right)\right) \in \mathscr{F}
$$

since $F_{4} \nsubseteq E_{1} \cup E_{3}$ and Case 1 has been proved. Thus

$$
\left(F_{1} \cap E_{1} \cap E_{3}\right) \cup\left(F_{2} \cap\left(E_{2} \cup E_{4}\right)\right) \in \mathscr{F}
$$

Case 2b. No member of $\mathscr{F}$ meets both $E_{1} \cap E_{3}$ and $E_{2} \cap E_{4}$. Then, since $H$ is non-separable, there exists $F_{3} \in \mathscr{F}$ meeting both $E_{1} \cup E_{3}$ and $E_{2} \cap E_{4}$.

Case $2 \mathrm{~b}(\mathrm{i}) . F_{3}$ meets $E_{1} \cap E_{4}$. If $F_{4} \in \mathscr{F}$ meets both $E_{1} \cap E_{3}$ and $E_{2} \cap E_{3}$, then $\left(F_{3} \cap E_{2}\right) \cup\left(F_{4} \cap E_{1}\right) \in \mathscr{F}$ meets both $E_{1} \cap E_{3}$ and $E_{2} \cap E_{4}$, a contradiction. Thus every member of $\mathscr{F}$ meeting $E_{1} \cap E_{3}$ is contained in $E_{1}$. But then

$$
\begin{aligned}
& \left(F_{1} \cap E_{1} \cap E_{3}\right) \cup\left(F_{2} \cap\left(E_{2} \cup E_{4}\right)\right) \\
& \quad=\left(F_{1} \cap E_{3}\right) \cup\left(F_{2} \cap E_{4}\right) \in \mathscr{F} .
\end{aligned}
$$

Case 2 b (ii). $F_{3}$ meets $E_{2} \cap E_{3}$. This case is similar to Case 2 b (i).
The proof is complete.
Theorem 15. ( $\mathscr{H}, E, \rightarrow)$ has the transitivity property.
Proof. Suppose that $\{\{x, y\}, E \backslash\{x, y\}\}$ and $\{\{y, z\}, E \backslash\{y, z\}\}$ are splits of $H=(E, \mathscr{F}) \in \mathscr{H}$ for distinct cells $x, y, z$ of $H$. We must show that $\{\{x, z\}, E \backslash\{x, z\}\}$ is a split of $H$. If $|E|>4$, then $\{\{x, y, z\}, E \backslash\{x, y, z\}\}$ is a split of $H$ by Theorem 14. Thus $H \rightarrow\left\{H_{1}, H_{2}\right\}$, where $H_{1}=$ $\left(\{x, y, z, w\}, \mathscr{F}_{1}\right)$ and by Proposition $5,\{\{x, y\},\{z, w\}\}$ and $\{\{y, z\}$, $\{x, w\}\}$ are splits of $H_{1}$, and $\{\{x, z\}, E \backslash\{x, z\}\}$ is a split of $H$ if and only if $\{\{x, z\},\{y, w\}\}$ is a split of $H_{1}$. Therefore, it is sufficient to prove the theorem in the case where $|E|=4$, say $E=\{x, y, z, w\}$.

Now let $F_{1}, F_{2}$ be arbitrary members of $\mathscr{F}$ meeting $\{x, z\}$ and $\{y, w\}$. We must show that

$$
F=\left(F_{1} \cap\{x, z\}\right) \cup\left(F_{2} \cap\{y, w\}\right) \in \mathscr{F} .
$$

If $F_{1} \cap\{y, w\}=F_{2} \cap\{y, w\}$, then $F=F_{1} \in \mathscr{F}$. Similarly, if $F_{1} \cap$ $\{x, z\}=F_{2} \cap\{x, z\}$, there is no problem. Also, for example, the case $y \in F_{1}, w \notin F_{1}$ is symmetrical to the case $y \notin F_{1}, w \in F_{1}$. By such considerations, we can assume that $x, y \in F_{1}$ and $z, w \in F_{2}$. For $i=1,2$ and $c=x, y, z, w$, let $F_{i}(c)=\{c\}$ if $c \in F_{i}$, and $\emptyset$ otherwise. Thus

$$
F_{1}=\{x, y\} \cup F_{1}(z) \cup F_{1}(w) ; F_{2}=F_{2}(x) \cup F_{2}(y) \cup\{z, w\}
$$

$\{\{y, z\},\{x, w\}\}$ is a split, and so

$$
\left(F_{1} \cap\{y, z\}\right) \cup\left(F_{2} \cap\{x, w\}\right) \in \mathscr{F} \text { and }
$$

$$
\left(F_{2} \cap\{y, z\} \cup\left(F_{1} \cap\{x, w\}\right) \in \mathscr{F} ;\right.
$$

that is

$$
\begin{aligned}
& F_{3}=\{y, w\} \cup F_{1}(z) \cup F_{2}(x) \in \mathscr{F} \text { and } \\
& F_{4}=\{x, z\} \cup F_{1}(w) \cup F_{2}(y) \in \mathscr{F} .
\end{aligned}
$$

But $\{\{x, y\},\{z, w\}\}$ is a split, and so $\left(F_{4} \cap\{x, y\}\right) \cup\left(F_{3} \cap\{z, w\}\right) \in \mathscr{F}$; that is,

$$
\{x, w\} \cup F_{1}(z) \cup F_{2}(y) \in \mathscr{F}
$$

But this is $\left(F_{1} \cap\{x, z\}\right) \cup\left(F_{2} \cap\{y, w\}\right)$, as required.
We have completed the proof of Theorem 7. We now prove Theorem 9, and thus Theorems 10 and 11. A set family $H=(E, \mathscr{F})$ is a throng if, whenever $E \supseteq F_{2} \supseteq F_{1} \in \mathscr{F}$, then $F_{2} \in \mathscr{F}$.

Proof of Theorem 9. It is straightforward to verify that bonds, stars, and superstars are brittle and non-separable. Now suppose that $H=$ $(E, \mathscr{F}) \in \mathscr{H}$ is brittle. We first prove:

Claim 1. If $F_{1} \subseteq F_{2} \subseteq F_{3}$ and $F_{1}, F_{3} \in \mathscr{F}$, then $F_{2} \in \mathscr{F}$.
Choose $x \in F_{1}$, and put $E_{1}=F_{2}-x, E_{2}=E \backslash E_{1}$. Then $\left|E_{1}\right| \geqq 2 \leqq$ $\left|E_{2}\right|$, and so $\left\{E_{1}, E_{2}\right\}$ is a split. But $F_{1}, F_{3}$ meet $E_{1}, E_{2}$, and so

$$
\left(F_{3} \cap E_{1}\right) \cup\left(F_{1} \cap E_{2}\right)=F_{2} \in \mathscr{F}
$$

proving Claim 1. We now prove:
Claim 2. $H$ is a throng or a bond or a star.
Let $F_{1}$ be a maximal member of $\mathscr{F}$. If $H$ is not a throng, then by Claim $1 F_{1} \neq E$, and there exists a maximal member $F_{2}$ of $\mathscr{F}$ distinct from $F_{1}$. By the non-separability of $H$, we may choose $F_{2}$ so that $F_{1}$ meets $F_{2}$. Suppose that $\left|F_{1} \cap F_{2}\right| \geqq 2$, and choose $x \in F_{1} \cap F_{2}$. Let $E_{1}=$ $\left(F_{1} \backslash F_{2}\right)+x$, and let $E_{2}=E \backslash E_{1}$. Then $F_{1}, F_{2}$ meet $E_{1}, E_{2}$ so

$$
\left(F_{1} \cap E_{1}\right) \cup\left(F_{2} \cap E_{2}\right)=F_{1} \cup F_{2} \in \mathscr{F}
$$

which contradicts the maximality of $F_{1}$ and $F_{2}$. Thus we may assume that $F_{1} \cap F_{2}=\{x\}$. Now suppose that $\left|F_{1}\right| \geqq 3$. Choose $y \in F_{1}-x$ and let $E_{1}=\{x, y\}$ and $E_{2}=E \backslash E_{1}$. Then $F_{1}, F_{2}$ meet $E_{1}, E_{2}$ so

$$
\left(F_{1} \cap E_{1}\right) \cup\left(F_{2} \cap E_{2}\right)=F_{2}+y \in \mathscr{F}
$$

contradicting the maximality of $F_{2}$. Thus $\left|F_{1}\right|=2$. It follows that all maximal members of $\mathscr{F}$, and thus all members of $\mathscr{F}$, have cardinality 2. Now we may assume that $E$ is the vertex-set of a simple, connected graph $G$ and that $\mathscr{F}$ consists of adjacent pairs of elements of $E$. Suppose that, for distinct elements $x, y, z, w$ of $E$, we have $\{w, x\},\{x, y\}$, $\{y, z\} \in \mathscr{F}$. Then $\{w, x\},\{y, z\}$ meet $\{w, z\}, E \backslash\{w, z\}$, so

$$
(\{w, x\} \cap\{w, z\}) \cup(\{y, z\} \cap E \backslash\{w, z\}))=\{w, y\} \in \mathscr{F} .
$$

It follows that vertices of $G$ are at distance 2 only if they have degree 1 . Since $G$ is connected, no distances greater than 2 can occur. Thus every vertex of $G$ has degree 1 or is adjacent to every other vertex. If all the vertices are of the second type, $G$ is complete and so $H$ is a bond. If there exist vertices of degree 1 , then there can be only one vertex which is adjacent to all of them. Thus $H$ is a star. This completes the proof of Claim 2.

Let $\mathscr{C}$ be the set of minimal members of $\mathscr{F}$. We now prove:
Claim 3. If $H$ is a throng, then where $A=\cup\{C: C \in \mathscr{C}\},|A| \geqq 2$ and $H^{\prime}=(A, \mathscr{C})$ is a non-separable brittle clutter.

Choose $E_{1} \subseteq E$ and suppose that $C_{1}, C_{2} \in \mathscr{C}$ meet $E_{1}$ and $E_{2}=E \backslash E_{1}$. Then

$$
F=\left(C_{1} \cap E_{1}\right) \cup\left(C_{2} \cap E_{2}\right) \in \mathscr{F} .
$$

Suppose $F \notin \mathscr{C}$. Then since $H$ is a throng, there exists $x \in F$ such that $F-x \in \mathscr{F}$. Assume $x \in E_{1}$. (The case $x \in E_{2}$ is similar.) If $C_{1} \cap$ $E_{1}=\{x\}$, then $F-x$ is a proper subset of $C_{2}$, which is not possible. We conclude that $F-x$ meets $E_{1}$ and $E_{2}$. Thus

$$
F_{1}=\left((F-x) \cap E_{1}\right) \cup\left(C_{1} \cap E_{2}\right) \in \mathscr{F} .
$$

But this is a contradiction, since $F_{1}$ is a proper subset of $C_{1}$. Thus $F \in \mathscr{C}$. Now suppose there exist distinct elementary separators $S_{1}, S_{2}$ of $(E, \mathscr{C})$ with $\left|S_{1}\right| \geqq 2 \leqq\left|S_{2}\right|$. Choose $D_{1} \subseteq E$ such that $S_{1}$ and $S_{2}$ meet both $D_{1}$ and $D_{2}=E \backslash D_{1}$, and choose $C_{i} \in \mathscr{C}$ such that $C_{i} \subseteq S_{i}$ and meets $D_{1}, D_{2}$ for $i=1$ and 2 . Then

$$
C=\left(C_{1} \cap D_{1}\right) \cup\left(C_{2} \cap D_{2}\right) \in \mathscr{C} .
$$

But $C$ meets $S_{1}, S_{2}$ which contradicts the fact that $S_{1}, S_{2}$ are elementary separators of $(E, \mathscr{C})$. Thus there is at most one elementary separator of ( $E, \mathscr{C}$ ) having cardinality greater than 1 . Also, unless $|E| \leqq 1$, when the theorem is obvious, there is one such elementary separator, say $A$. Since $H$ has no loops, $\mathscr{C}=\mathscr{C} \backslash(E \backslash A)$. Thus $H^{\prime}=(A, \mathscr{C})$ is a non-separable brittle clutter, and Claim 3 is proved.

Now $H^{\prime}$ must be a bond or a star or a polygon by Claim 2. If $H^{\prime}$ is a bond with $A=E$, then $H$ is a 0 -superstar. If $H^{\prime}$ is a star with $A=E$, then $H$ is a 1 -superstar. If $H^{\prime}$ is a polygon and $|A|=k \geqq 2$, then $H$ is a $k$-superstar. Thus we must eliminate the possibility that $H^{\prime}$ is a star or a bond (not a polygon) with $A \neq E$. If this happens, we have $|A| \geqq 3$, since a bond or star $H^{\prime}$ with $|A|=2$ is also a polygon. There exist distinct elements $x, y, z$ of $A$ with $\{x, y\},\{y, z\} \in \mathscr{C}$. Choose $w \in E \backslash A$ and $E_{1} \subseteq E$ such that $x, y \in E_{1}$ and $w, z \in E_{2}=E \backslash E_{1}$. By the definition of

$$
\begin{aligned}
& \mathscr{C},\{x, y, w\} \in \mathscr{F} . \text { Also }\{x, y, w),\{y, z\} \text { meet } E_{1}, E_{2} \text { so } \\
& \quad\left(\{x, y, w\} \cap E_{2}\right) \cup\left(\{y, z\} \cap E_{1}\right)=\{y, w\} \in \mathscr{F} .
\end{aligned}
$$

But this is a contradiction, since $\{y, w\}$ does not contain a member of $\mathscr{C}$. The proof of the theorem is complete.

The proof of the above theorem, revealing as it does the close connection between a brittle throng and its clutter of minimal members, might lead one to suspect that the clutter of minimal members of an arbitrary set family would yield a great deal of information about the possible decompositions of the set family. This does not seem to be the case, and indeed the authors believe that the decomposition theory for arbitrary set families cannot easily be derived from the corresponding clutter theory.
6. Matroid decomposition. A set family $M=(E, \mathscr{F})$ is said to be a matroid if $\mathscr{F}$ is a clutter, and the set
$\mathscr{I}=\{J \subseteq E: J$ contains no member of $\mathscr{F}\}$
(called the set of independent sets of $M$ ) satisfies:
I1. For all $A \subseteq E$, any two maximal independent subsets of $A$ have the same cardinality.

A maximal independent subset of $A$ is called a basis of $A$, and its cardinality (which depends only on $A$ ) is called the rank, $r(A)$, of $A$. Since we often deal with more than one matroid, we occasionally prefix terms by the name of the appropriate matroid; for example, " $M$-independent." Similarly, the rank function of matroid $M$ (or $M_{i}$, or $M^{\prime}$ ) will be denoted by $r$ (or $r_{i}$, or $r^{\prime}$ ). The set family which we have associated with the matroid $M$ has as its members the minimal dependent (non-independent) sets of $M$, called the circuits of $M$. For this special case, the set-family composition was formulated by Minty [18]. We will define any matroid terminology used here, but we occasionally use well-known, elementary facts without proving them; a good matroid-theory reference is [27].

The notion of separability which we have used for set families is well known for matroids. A standard result concerning it says that $A \subseteq E$ is a separator of $M$ if and only if $r(A)+r(E \backslash A)=r(E)$. Tutte [24] has defined a partition $\left\{E_{1}, E_{2}\right\}$ of $M$ to be a $k$-separation of $M$, for $k$ a positive integer, if $\left|E_{1}\right| \geqq k \leqq\left|E_{2}\right|$ and $r\left(E_{1}\right)+r\left(E_{2}\right) \leqq r(E)+k-1$. For $n$ a positive integer, the matroid $M$ is said to be $n$-connected if $M$ has no $k$-separation for any positive integer $k<n$. Thus, in particular, every matroid is 1 -connected, and a matroid is 2 -connected if and only if it is non-separable. We show that a matroid $M \in \mathscr{H}$ is prime if and only if it
is 3 -connected. This result has also been obtained in [6] and [19]; related work is that of [8] and [21].

Theorem 16. Let $M \in \mathscr{H}$ be a matroid, and let $\left\{E_{1}, E_{2}\right\}$ be a partition of $E=E(M)$. Then $\left\{E_{1}, E_{2}\right\}$ is a split of $M$ if and only if $\left\{E_{1}, E_{2}\right\}$ is a 2-separation of $M$.

Lemma 12. Let $\left\{E_{1}, E_{2}\right\}$ be a 2-separation of the non-separable matroid $M$ and let $B_{1}$ be a basis of $E_{1}$. Then $B_{1}$ contains a set $D$ such that every circuit $C$ having $\emptyset \neq C \cap B_{1}=C \cap E_{1}$ satisfies $C \cap B_{1}=D$.

Proof. Where $B_{2}$ is a basis of $E_{2}, B_{1} \cup B_{2}$ contains a unique circuit $C$, because $B_{1} \cup B_{2}=B+x$ for some basis $B$ of $E$ and $x \notin B$. It will be enough to show that, for every choice of $B_{2}, C \cap B_{1}$ is the same set $D$, since every circuit of the kind described in the lemma arises in this way. Moreover, since $b \in B_{2} \backslash B_{2}{ }^{\prime}$ for bases $B_{2}, B_{2}{ }^{\prime}$ of $E_{2}$ implies that there exists $b^{\prime} \in B_{2}{ }^{\prime} \backslash B_{2}$ such that $B_{2}-b+b^{\prime}$ is also a basis of $E_{2}$, it will be enough to consider bases $B_{2}$ and $B_{2}{ }^{\prime}=B_{2}-b+b^{\prime}$ of $E_{2}$. Let $C, C^{\prime}$ be the circuit contained in $B_{1} \cup B_{2}, B_{1} \cup B_{2}{ }^{\prime}$, respectively. There is a circuit $F \subseteq B_{2}+b^{\prime}$, and $b^{\prime} \in F$. For any $x \in\left(C^{\prime} \cap B_{1}\right) \backslash C$, there exists a circuit $G$ such that $x \in G \subseteq\left(C^{\prime} \cup F\right)-b^{\prime}$. But $G \subseteq B_{1} \cup B_{2}$, so $G=C$; that is, $C \supseteq C^{\prime} \cap B_{1}$. Similarly, $C^{\prime} \supseteq C \cap B_{1}$, and the proof is finished.

Proof of Theorem 16. Suppose that $\left\{E_{1}, E_{2}\right\}$ is a split of $M$. Let $B_{1}$ be a basis of $E_{1}$. Extend $B_{1}$ to a basis $B$ of $E$, and extend $B \cap E_{2}$ to a basis $B_{2}$ of $E_{2}$. We must show that $\left|B_{2} \backslash B\right|=1$. If not, then there exist distinct elements $x, y$ of $B_{2} \backslash B$. There exist circuits $C_{x}, C_{y}$ such that $x \in C_{x} \subseteq$ $B+x, y \in C_{y} \subseteq B+y$, and both $C_{x}$ and $C_{y}$ must meet $E_{1}$. Then

$$
C_{x}^{\prime}=\left(C_{x} \cap E_{1}\right) \cup\left(C_{y} \cap E_{2}\right)
$$

is a circuit and there exists $z \in C_{x}{ }^{\prime} \cap C_{y}$. Thus there exists a circuit $C$ such that

$$
x \in C \subseteq\left(C_{x}^{\prime} \cup C_{y}\right)-z
$$

Now $C$ must meet $E_{1}$, since otherwise $C \subseteq B_{2}$, but then

$$
C^{\prime}=\left(C_{x}{ }^{\prime} \cap E_{1}\right) \cup\left(C \cap E_{2}\right)
$$

is a circuit properly contained in $C_{x}{ }^{\prime}$, a contradiction. Therefore, $\left\{E_{1}, E_{2}\right\}$ is a 2 -separation of $M$.

Now suppose that $\left\{E_{1}, E_{2}\right\}$ is a 2 -separation of $M$, and let $C, C^{\prime}$ be circuits of $M$ meeting $E_{1}$ and $E_{2}$. Extend $C \cap E_{1}$ to a basis $B_{1}$ of $E_{1}$, and extend $C^{\prime} \cap E_{2}$ to a basis $B_{2}$ of $E_{2}$. It follows from Lemma 12 that every circuit $F$ meeting $E_{1}$ and $E_{2}$ such that $F \cap E_{1} \subseteq B_{1}$ satisfies $F \cap E_{1}=C \cap E_{1}$; similarly, if $F \cap E_{2} \subseteq B_{2}$, then $F \cap E_{2}=C^{\prime} \cap E_{2}$. But $B_{1} \cup B_{2}$ contains such a circuit $F$, so $F=\left(C \cap E_{1}\right) \cup\left(C^{\prime} \cap E_{2}\right)$.

The decomposition (and composition) for matroids can be related to certain standard matroid constructions, as follows. Given a matroid $M=(E, \mathscr{F})$, the set family $M \backslash A$ is also a matroid, obtained from $M$ by deleting $A$. The set family $M / A$ can also be proved to be a matroid, obtained from $M$ by contracting $A ; M / A$ is defined to have as circuits the minimal sets $C \subseteq E \backslash A$ such that $C \cup A$ contains an $M$-circuit meeting $C$. Where $e \in E$, we will abbreviate $M \backslash\{e\}$ to $M \backslash e$ and $M /\{e\}$ to $M / e$. A set $A \subseteq E$ is a series set of $M$ if every $M$-circuit $C$ satisfies $C \supseteq A$ or $C \cap A=\emptyset$. A series contraction is a contraction of a proper subset of a series set. A series minor of $M$ is a matroid obtained from $M$ by a deletion followed by a sequence of series contractions. Finally, the sum [11], $M_{1}+M_{2}$, of matroids $M_{1}$ and $M_{2}$ is a matroid satisfying $E\left(M_{1}+M_{2}\right)=E\left(M_{1}\right) \cup E\left(M_{2}\right)$, a set being $\left(M_{1}+M_{2}\right)$-independent if and only if it can be expressed as a union of an $M_{1}$-independent set with an $M_{2}$-independent set. (Note that $E\left(M_{1}\right), E\left(M_{2}\right)$ need not be disjoint.)

Theorem 17. Let $M=(E, \mathscr{F}) \in \mathscr{H}$ be a matroid, and let $\left\{M_{1}, M_{2}\right\}$ be the simple decomposition of $M$ associated with the split $\left\{E_{1}, E_{2}\right\}$ of $M$ and the marker e. Then $M_{1}$ and $M_{2}$ are isomorphic to series minors of $M$ (and therefore are matroids). Moreover, $M=\left(M_{1}+M_{2}\right) / e$.

Proof. Since $M \in \mathscr{H}$, we can choose an $M$-circuit $C$ meeting $E_{1}$ and $E_{2}$. Let $D=E_{2} \cap C$, and choose $x \in D$. Let

$$
M_{1}^{\prime}=\left(M \backslash\left(E_{2} \backslash D\right)\right) /(D-x)
$$

It is easy to see that $D$ is a series set of $M \backslash\left(E_{2} \backslash D\right)$, and thus $M_{1}{ }^{\prime}$ is a series minor of $M$. We claim that $M_{1}$ can be obtained from $M_{1}{ }^{\prime}$ by replacing $x$ by $e$. To prove this, it will be enough to show that the set of circuits of $M_{1}{ }^{\prime}$ is

$$
\begin{aligned}
&\left\{C \subseteq E_{1}: C \text { an } M \text {-circuit }\right\} \cup\left\{C \cap E_{1}+x:\right. \\
&\left.C \text { an } M \text {-circuit meeting } E_{1} \text { and } E_{2}\right\} .
\end{aligned}
$$

But this follows from the fact that $D$ is a series set of $M \backslash\left(E_{2} \backslash D\right)$. Therefore, $M_{1}$ is isomorphic to a series minor of $M$, and similarly for $M_{2}$.
To prove the second part of the theorem, observe that the set of ( $M_{1}+M_{2}$ )-circuits is

$$
\begin{aligned}
& \left\{C \subseteq E_{1}: C \text { an } M_{1} \text {-circuit }\right\} \cup\left\{C \subseteq E_{2}: C \text { an } M_{2} \text {-circuit }\right\} \\
& \cup\left\{C_{1} \cup C_{2}: e \in C_{1} \cap C_{2}, C_{i} \text { an } M_{i} \text {-circuit for } i=1 \text { and } 2\right\} .
\end{aligned}
$$

It follows that the circuits of $\left(M_{1}+M_{2}\right) / e$ are precisely the circuits of $M$.
The only brittle set families which are matroids are the bonds and polygons. Therefore, we derive from Theorem 10 the following unique decomposition theorem for matroids.

Theorem 18. Every non-separable matroid has a unique minimal decomposition, each of whose members are bonds, polygons, or 3-connected matroids.

We point out that the proof we have given for Theorem 18 is far from being the shortest possible. In particular, Theorem 18 can be derived quite simply from Theorem 5, since the intersection and transitivity properties are extremely easy to prove for matroids, as is a characterization of the brittle matroids. In fact, even the work of Section 3 can be avoided; in [10], Theorem 18 is derived from Theorem 3. A main tool in that derivation is a matroid-theoretic characterization of the good splits of a matroid, which we give here without proof.

Theorem 19. Let $M$ be a non-separable matroid. The 2-separation $\left\{E_{1}, E_{2}\right\}$ of $M$ is a good split of $M$ if and only if at least one of $M \backslash E_{1}$, $M \backslash E_{2}$ is non-separable, and at least one of $M / E_{1}, M / E_{2}$ is non-separable.

As before, let us call the decomposition whose uniqueness is asserted in Theorem 18, the standard decomposition of $M$. The next result (whose proof we omit; see [10]) extends some basic connectivity results of Tutte. He proves [24] that, where $e$ is a cell of a non-separable matroid $M$, at least one of $M \backslash e, M / e$ is non-separable; moreover, if $M$ is 3-connected, and has at least 4 cells, both $M \backslash e$ and $M / e$ are non-separable.

Theorem 20. Let e be a cell of a non-separable matroid $M$ having at least 3 cells, and let $M^{\prime}$ be the member of the standard decomposition of $M$ which has e as a cell. Then
(a) $M \backslash e$ is separable if and only if $M^{\prime}$ is a polygon;
(b) $M / e$ is separable if and only if $M^{\prime}$ is a bond.

The fact that the members of any decomposition of a matroid $M$ are series minors of $M$ implies that many important classes of matroids are "closed under decomposition". These include polygon matroids, matroids linear over a given field, and transversal matroids. (Similarly, many well-known classes are "closed under composition".) In the case of polygon matroids, Tutte [24] has shown that a connected graph $G$ is 3 -connected if and only if $P M(G)$ is a 3 -connected matroid. Therefore, the polygon matroids of the members of the standard (graph) decomposition $D$ of $G$ constitute a (matroid) decomposition $D^{\prime}$ of $P M(G)$, each of whose members is a bond, a polygon, or a 3 -connected matroid, and $D^{\prime}$ is minimal with this property. Therefore, $D^{\prime}$ is the standard decomposition of $P M(G)$. In spite of this close relationship between the two standard decompositions, Theorem 1 applied to $G$ and Theorem 18 applied to $P M(G)$ are quite different uniqueness results.

Perhaps the most important aspect of matroid decomposition as a special case of the set-family theory is that matroids constitute the only
identifiable class of set families for which efficient methods of constructing the standard decomposition are known. These methods are based on algorithms for finding matroid $k$-separations. In [10] a particularly simple and efficient algorithm for separability was described; this method, for the case of binary matroids, is implicit in the papers of Tutte. A good algorithm for finding a $k$-separation, if one exists, for any fixed $k$, based on the matroid partition [11] (or matroid intersection [12]) algorithm, was described in [10]. More recently, an efficient recursive algorithm for testing for 3 -connectivity was discovered [7].

The expression for $M$ in terms of $M_{1}$ and $M_{2}$ in Theorem 17 suggests a more general composition for matroids: forming the matroid $\left(M_{1}+M_{2}\right) /\left(E_{1} \cap E_{2}\right)$, where $M_{i}$ is on $E_{i}$ for $i=1$ and 2 . This composition is investigated in [10] (and also in [19]); many, but not all, of the attractive properties of the cases $\left|E_{1} \cap E_{2}\right|=0$ or 1 extend to the more general composition. Another topic of investigation [10] has been a decomposition theory for systems of homogeneous linear equations, for which " $k$-decomposability" is equivalent to $(k+1)$-separability of the associated matroid. In particular, there is a unique decomposition theory for the case in which simple decompositions are pairs of systems having exactly one common variable, and this theory is closely related to the matroid decomposition theory for the matroid associated with a linear system. We do not treat the linear system decomposition here, partly because it does not satisfy the decomposition frame axioms. The linear system theory, the generalized matroid decomposition, and the matroid connectivity algorithms, will be described in another paper.
7. The substitution decomposition. In this section we investigate a notion of decomposition, "substitution decomposition", which has been studied by several other authors. The theory associated with this decomposition is shown to be a special case of the clutter decomposition theory of Section 5. By applying the general set family theory, we generalize the substitution decomposition theory for clutters to arbitrary set families.

Let $H_{1}=\left(E_{1}+e, \mathscr{F}_{1}\right)$ and $H_{2}=\left(E_{2}, \mathscr{F}_{2}\right)$ be set families such that $e \notin E,\left\{E_{1}, E_{2}\right\}$ is a partition of $E$, and $\left|E_{1}\right| \geqq 1,\left|E_{2}\right| \geqq 2$. We define $H_{1}\left[H_{2} ; e\right]$ to be the set family $H=(E, \mathscr{F})$, where

$$
\widetilde{F}=(\mathscr{F} \backslash e) \cup\left\{F_{1} \cup F_{2}-e: e \in F_{1} \in \mathscr{F}_{1}, F_{2} \in \mathscr{F}_{2}\right\}
$$

We call $H_{1}\left[H_{2} ; e\right]$ a substitution composition. Despite certain similarities between this composition and the set family composition described in Section 5, there are the following obvious differences. Here the two sets of cells are disjoint, the composition is not commutative, and the composition depends on the choice of a "special" cell $e$ of $H_{1}$.

The substitution composition for the special case in which the set families are clutters has been studied previously in several contexts:

Boolean functions ([1], [5]), simple games ([20]), and clutters ([4]). As an illustration of the substitution composition, we briefly outline the connection with Boolean functions. Let $E=\{1,2, \ldots, n\}$. A function $f$ from $\{0,1\}^{E}$ to $\{0,1\}$ is monotone if

$$
\begin{aligned}
f\left(x_{1}, x_{2}, \ldots, x_{n}\right) \leqq f\left(y_{1}, y_{2}, \ldots, y_{n}\right) \text { whenever } x_{i} \leqq y_{i} \\
\text { for } i=1,2, \ldots, n .
\end{aligned}
$$

We assume that $f(0,0, \ldots, 0)=0$. If $A \subseteq E, x^{A} \in\{0,1\}^{E}$ denotes ( $x_{i}: i \in E$ ) where $x_{i}=1$ if and only if $i \in A$. Associated with the monotone function $f$ is the clutter $H=(E, \mathscr{F})$ defined by: $F \in \mathscr{F}$ if and only if $F$ is minimal such that $f\left(x^{F}\right)=1$. It is clear that $f$ is also recoverable from $H$. In these circumstances we denote $H$ by $c(f)$. The following result, whose proof is straightforward, can be found in [3].

Proposition 3. Let $E_{1}=\{1,2, \ldots k\}$ and $E_{2}=\{k+1, \ldots, n\}$, where $k \geqq 1$ and $n-k \geqq 2$. Let e be an element not in $E$. For monotone functions $g$ from $\{0,1\}^{E_{1}+e}$ to $\{0,1\}$ and $h$ from $\{0,1\}^{E_{2}}$ to $\{0,1\}$, we have

$$
f\left(x_{1}, x_{2}, \ldots, x_{n}\right)=g\left(x_{1}, x_{2}, \ldots, x_{k}, h\left(x_{k+1}, \ldots, x_{n}\right)\right)
$$

for every $\left(x_{1}, \ldots, x_{n}\right) \in\{0,1\}^{E}$ if and only if $c(f)=c(g)[c(h) ; e]$.
If $\left\{E_{1}, E_{2}\right\}$ is a partition of $E$, we say that $E_{2}$ is a committee of $H=(E, \mathscr{F})$ if $\left|E_{1}\right| \geqq 1,\left|E_{2}\right| \geqq 2$ and, whenever $F_{1}, F_{2} \in \mathscr{F}$ both meet $E_{2}$,

$$
\left(F_{1} \cap E_{1}\right) \cup\left(F_{2} \cap E_{2}\right) \in \mathscr{F} .
$$

A cell $e$ of $H$ is an isthmus of $H$ if there exists no member $F$ of $\mathscr{F}$ containing $e$.

Proposition 4. Let $H=(E, \mathscr{F})$ be a set family and let $\left\{E_{1}, E_{2}\right\}$ be a partition of $E$ with $\left|E_{1}\right| \geqq 1,\left|E_{2}\right| \geqq 2$. There exist set families $H_{1}=\left(E_{1}+e, \mathscr{F}_{1}\right)$ and $H_{2}=\left(E_{2}, \mathscr{F}_{2}\right)$, where $e \notin E$, such that $H=H_{1}\left[H_{2} ; e\right]$ if and only if $E_{2}$ is a committee of $H$. Moreover, if $H$ has no isthmus, then $H_{1}$ and $H_{2}$ are uniquely determined, by

$$
\begin{aligned}
& \mathscr{F}_{1}=\left(\mathscr{F} \backslash E_{2}\right) \cup\left\{F \cap E_{1}+e: F \in \mathscr{F} \text { meets } E_{2}\right\} \text { and } \\
& \mathscr{F}_{2}=\left\{F \cap E_{2}: F \in \mathscr{F} \text { meets } E_{2}\right\},
\end{aligned}
$$

and neither $H_{1}$ nor $H_{2}$ has an isthmus.
Proof. It is clear from the definitions that, if $H=H_{1}\left[H_{2} ; e\right]$, then $E_{2}$ is a committee of $H$.

If $E_{2}$ is a committee of $H$, define $\mathscr{F}_{1}$ to be

$$
\left(\mathscr{F} \backslash E_{2}\right) \cup\left\{F \cap E_{1}+e: F \in \mathscr{F}, F \text { meets } E_{2}\right\}
$$

and $\mathscr{F}_{2}$ to be

$$
\left\{F \cap E_{2}: F \in \mathscr{F}, F \text { meets } E_{2}\right\} .
$$

Then it is straightforward to check that, if $H_{1}=\left(E_{1}+e, \mathscr{F}_{1}\right)$ and $H_{2}=\left(E_{2}, \mathscr{F}_{2}\right)$, we have $H=H_{1}\left[H_{2} ; e\right]$.

Now suppose that $E_{2}$ is a committee of $H$ and $H=H_{1}{ }^{\prime}\left[H_{2}{ }^{\prime} ; e\right]$, where $H_{1}{ }^{\prime}=\left(E_{1}+e, \mathscr{F}_{1}{ }^{\prime}\right)$ and $H_{2}{ }^{\prime}=\left(E_{2}, \mathscr{F}_{2}{ }^{\prime}\right)$, and suppose that $H$ has no isthmus. Then

$$
\left\{F \in \mathscr{F}: F \text { meets } E_{2}\right\} \neq \emptyset .
$$

Thus $\mathscr{F}_{2}{ }^{\prime} \neq \emptyset$.
Let $F_{1}{ }^{\prime} \in \mathscr{F}{ }_{1}{ }^{\prime}$. If $e \notin F_{1}{ }^{\prime}$, then $F_{1}{ }^{\prime} \in \mathscr{F} \backslash E_{2}$, so $F_{1}{ }^{\prime} \in \mathscr{F}{ }_{1}$. If $e \in F_{1}{ }^{\prime}$, choose $F_{2}{ }^{\prime} \in \mathscr{F}_{2}{ }^{\prime}$. Then $F=F_{1}{ }^{\prime} \cup F_{2}{ }^{\prime}-e \in \mathscr{F}, \quad$ and $F_{1}{ }^{\prime}=F \cap$ $E_{1}+e$, so $F_{1}{ }^{\prime} \in \mathscr{F}_{1}$. Thus $\mathscr{F}_{1}{ }^{\prime} \subseteq \mathscr{F}_{1}$.

Now suppose $F_{1} \in \mathscr{F}{ }_{1}{ }^{\prime}$. If $e \notin \bar{F}_{1}{ }^{\prime}$, then $F_{1} \in \mathscr{F} \backslash E_{2}$. Thus $F_{1} \in \mathscr{F}{ }_{1} \backslash e$, so $F_{1} \in \mathscr{F}_{1}{ }^{\prime}$. If $e \in F_{1}$, then $F_{1}=F \cap E_{1}+e$ for some $F \in \mathscr{F}$ such that $F$ meets $E_{2}$. Then $F=F_{1}{ }^{\prime} \cup F_{2}{ }^{\prime}-e$, where $e \in F_{1}{ }^{\prime} \in \mathscr{F}_{1}{ }^{\prime}$ and $F_{2}{ }^{\prime} \in \mathscr{F}_{2}{ }^{\prime}$. But then $F_{1}=F_{1}{ }^{\prime}$, so $F_{1} \in \mathscr{F}_{1}{ }^{\prime}$. It follows that $\mathscr{F}_{1}{ }^{\prime}=\mathscr{F}_{1}$.

Now suppose that $F_{2}{ }^{\prime} \in \mathscr{F}_{2}{ }^{\prime}$. Since some $F \in \mathscr{F}$ meets $E_{2}, e$ is not an isthmus of $H_{1}{ }^{\prime}$. Therefore, there exists $F_{1}{ }^{\prime} \in \mathscr{F}{ }_{1}{ }^{\prime}$ with $e \in F_{1}{ }^{\prime}$. Then $F=F_{1}{ }^{\prime} \cup F_{2}{ }^{\prime}-e \in \mathscr{F}$, so $F_{2}{ }^{\prime}=F \cap E_{2} \in \mathscr{F}_{2}$. Thus $\mathscr{F}_{2}{ }^{\prime} \subseteq \mathscr{F}_{2}$.

Finally, suppose that $F_{2} \in \mathscr{F}_{2}$. Then $F_{2}=F \cap E_{2}$ for some $F \in \mathscr{F}$ such that $F$ meets $E_{2}$. Now $F=F_{1}{ }^{\prime} \cup F_{2}{ }^{\prime}-e$, where $e \in F_{1}{ }^{\prime} \in \mathscr{F}_{1}{ }^{\prime}$ and $F_{2}{ }^{\prime} \in \mathscr{F}_{2}{ }^{\prime}$. Then $F_{2}=F_{2}{ }^{\prime}$, so $F_{2} \in \mathscr{F}_{2}{ }^{\prime}$. It follows that $\mathscr{F}_{2}=\mathscr{F}_{2}{ }^{\prime}$.

Therefore, $H_{1}{ }^{\prime}=H_{1}$ and $H_{2}{ }^{\prime}=H_{2}$, provided $H$ has no isthmus. In this case, it is clear that $H_{1}, H_{2}$ also have no isthmus, and the proof is complete.

It is possible to give examples to show that the uniqueness of $H_{1}, H_{2}$ can fail if $H$ is allowed to have isthmuses. Therefore, set families considered in this section will not have isthmuses. We do not exclude loops, or require non-separability.

The substitution composition seems to be harder to work with than the composition studied (implicitly) in Section 5. In developing a decomposition theory based on this composition, it will be convenient to consider the objects being decomposed to be pairs ( $H, e$ ), where $H$ is a set family and $e$ is not a cell of $H$. We will see that this device makes the composition easier to handle by symmetrizing it.

Suppose that $H=H_{1}\left[H_{2} ; e_{2}\right]$. Let $e_{1}$ be an element which is not a cell of $H$ and is different from $e_{2}$. Then we say that $\left\{\left(H_{1}, e_{1}\right),\left(H_{2}, e_{2}\right)\right\}$ is a simple factorization of $\left(H, e_{1}\right)$. We say that $e_{2}$ is the marker of the simple factorization. A factorization of $(H, e)$ is defined inductively to be either $\{(H, e)\}$ or a set $D^{\prime}$ obtained from a factorization $D$ of $(H, e)$ by replacing a member $\left(H_{1}, e_{1}\right)$ of $D$ by the members of a simple factorization
of $\left(H_{1}, e_{1}\right)$, such that the marker of this simple factorization is neither a cell of $H$ nor an element $e^{\prime}$ such that $\left(H^{\prime}, e^{\prime}\right) \in D$ for some $H^{\prime}$. If $H^{\prime \prime}$ is obtained from $H$ by a (non-empty) sequence of operations of the kind described above, then $D^{\prime \prime}$ is said to be a (strict) refinement of $D$. A marker of a factorization $D$ of $(H, e)$ is an element $e^{\prime} \neq e$ such that $\left(H^{\prime}, e^{\prime}\right) \in D$ for some $H^{\prime}$. A component of $D$ is a set family $H^{\prime}$ such that $\left(H^{\prime}, e^{\prime}\right) \in D$ for some $e^{\prime}$. Clearly each marker of $D$ is a cell of exactly one component of $D$. A factorization $D$ is trivial if $|D|=1$.

Given a factorization $D$ of $(H, e)$, we may form a directed graph $G$ as follows: The vertices of $G$ are the components of $D$ and the edges are its markers; the marker $e^{\prime}$ is directed from $H_{1}$ to $H_{2}$, where $H_{1}$ is the component of $D$ such that $\left(H_{1}, e^{\prime}\right) \in D$ and $H_{2}$ is the component of $D$ of which $e^{\prime}$ is a cell. It is easy to see that $G$ is a tree with the property that every vertex but one is the tail of exactly one edge; the exceptional vertex is the component $H^{\prime}$ of $D$ such that $\left(H^{\prime}, e\right) \in D$. As for the tree associated with a decomposition in earlier sections, this directed tree provides a convenient way to visualize a factorization.

The terms "equivalent" and "minimal" are defined for factorizations in the same way as for decompositions. If $e$ is not a cell of the set family $H=(E, \mathscr{F})$, we define $e H$ to be the set family $(E+e, e \mathscr{F})$, where $e \mathscr{F}=\{F+e: F \in \mathscr{F}\}$. It is easy to see that $e H$ is loopless; also, $e H$ is non-separable if and only if $H$ has no isthmus. The following result links the present notion of factorization to that of decomposition, discussed in Section 5.

Theorem 21. $D$ is a factorization of $(H, e)$ if and only if $D^{\prime}=$ $\left\{e^{\prime} H^{\prime}:\left(H^{\prime}, e^{\prime}\right) \in D\right\}$ is a decomposition of $e H$.

Proof. The result is clearly true for $|D|=1$. From the definitions of factorization and decomposition, it is enough to prove the result for $|D|=2$. Suppose that $\left\{\left(H_{1}, e_{1}\right),\left(H_{2}, e_{2}\right)\right\}$ is a simple factorization of ( $H, e_{1}$ ), where

$$
H_{1}=\left(E_{1}+e_{2}, \mathscr{F}_{1}\right), H_{2}=\left(E_{2}, \mathscr{F}_{2}\right), \text { and } H=(E, \widetilde{F}) .
$$

Then

$$
\mathscr{F}=\left(\mathscr{F}_{1} \backslash e_{2}\right) \cup\left\{F_{1} \cup F_{2}-e_{2}: e_{2} \in F_{1} \in \mathscr{F}_{1}, F_{2} \in \mathscr{F}_{2}\right\} .
$$

Thus

$$
\begin{aligned}
e_{1} \mathscr{F} & =\left(e_{1}\left(\mathscr{F}_{1} \backslash e_{2}\right)\right) \cup\left\{F_{1} \cup F_{2}-e_{2}+e_{1}: e_{2} \in F_{1} \in \mathscr{F}_{1}, F_{2} \in \mathscr{F}_{2}\right\} \\
& =\left(e_{1}\left(\mathscr{F}_{1} \backslash e_{2}\right)\right) \cup\left\{F_{1} \cup F_{2}-e_{2} ; e_{2} \in F_{1} \in e_{1} \mathscr{F}_{1}, e_{2} \in F_{2} \in e_{2} \mathscr{F}_{2}\right\} .
\end{aligned}
$$

Since $\left(e_{2} \mathscr{F}_{2}\right) \backslash e_{2}=\emptyset$, it follows that $\left\{e_{1} H_{1}, e_{2} H_{2}\right\}$ is a simple decomposition of $e_{1} H^{\prime}$ with marker $e_{2}$.

Now suppose that $\left\{e_{1} H_{1}, e_{2} H_{2}\right\}$ is a simple decomposition of $e_{1} H$. Then,
since $\left(e_{2} \mathscr{F}_{2}\right) \backslash e_{2}=\emptyset$, we have

$$
\begin{aligned}
& e_{1} \mathscr{F}=\left(\left(e_{1} \mathscr{F}_{1}\right) \backslash e_{2}\right) \cup\left\{F_{1} \cup F_{2}-e_{2}: e_{2} \in F_{1} \in e_{1} \mathscr{F}_{1},\right. \\
&\left.e_{2} \in F_{2} \in e_{2} \mathscr{F}_{2}\right\} \\
&=\left(\left(e_{1} \mathscr{F}_{1}\right) \backslash e_{2}\right) \cup\left\{F_{1} \cup F_{2}-e_{2}+e_{1}: e_{2} \in F_{1} \in \mathscr{F}_{1}, F_{2} \in \mathscr{F}_{2}\right\} .
\end{aligned}
$$

Thus

$$
\widetilde{F}=\left(\mathscr{F}_{1} \backslash e_{2}\right) \cup\left\{F_{1} \cup F_{2}-e_{2}: e_{2} \in F_{1} \in \mathscr{F}_{1}, F_{2} \in \mathscr{F}_{2}\right\},
$$

so $\left\{\left(H_{1}, e_{1}\right),\left(H_{2}, e_{2}\right)\right\}$ is a simple factorization of $\left(H, e_{1}\right)$.
Corollary. If $\left\{E_{1}, E_{2}\right\}$ is a partition of $E$, then $E_{1}$ is a committee of $H$ if and only if $\left\{E_{1}+e, E_{2}\right\}$ is a split of $e H$.

A set family $H$ is irreducible if, whenever $e$ is not a cell of $H,(H, e)$ has no non-trivial factorization; $H=(E, \mathscr{F})$ is fragile if every subset $A \subset E$ such that $|A| \geqq 2$, is a committee of $H$. Applying Theorem 8 , we obtain the following result.

Theorem 22. Let $H$ be a set family having no isthmus and let $e$ be an element which is not a cell of $H$. Then $(H, e)$ has a unique minimal factorization, each of whose components is irreducible or fragile.

To characterize the fragile set families, we merely check the list (Theorem 9) of brittle set families; a set family $H$ is fragile if and only if $e H$ is brittle. Every brittle set family $H=(E, \mathscr{F})$, which has a cell $e \in E$ such that $e \in F$ for every $F \in \mathscr{F}$, gives rise to a fragile set family in this way. The ones that do not have this property are bonds and 0 -superstars. The fact that not all brittle families correspond in this way to fragile families is one indication that the theory of Section 5 is a proper generalization of the present theory. A set family $H=(E, \mathscr{F})$ is called a revised superstar if

$$
\mathscr{F}=\{F: A \subseteq F \subseteq E, F \neq \emptyset\} \text { for some } A \subseteq E ;
$$

$H$ is called a flower if

$$
\mathscr{F}=\{F: F \subseteq E,|F|=1\} .
$$

The following results are immediate consequences of Theorems 9 and 10.
THEOREM 23. A set family having no isthmus is fragile if and only if it is a flower or a revised superstar.

Theorem 24. Let $H$ be a set family having no isthmus, and $e$ be an element which is not a cell of $H$. Then $(H, e)$ has a unique minimal factorization $D$ such that each component of $D$ is irreducible, a flower, or a revised superstar.

Just as in Section 5, we can restrict the theory to clutters. It is easy to see that the components of a factorization of a clutter are themselves clutters. Notice that the fragile clutters are the flowers and the polygons. Thus we have the following result.

Theorem 25. Let $H$ be a clutter having no isthmus and let $e$ be an element which is not a cell of $H$. Then $(H, e)$ has a unique minimal factorization $D$ such that each component of $D$ is an irreducible clutter, a flower or a polygon.

Theorem 25 could be obtained from Theorem 11 in exactly the same way that Theorem 24 was obtained from Theorem 10 . Thus Theorem 10 may be said to generalize Theorem 25 in two different directions.

In view of Proposition 3, Theorem 25 yields a unique decomposition theorem for monotone Boolean functions. The fragile clutters, namely the polygons and flowers, have associated Boolean functions which are particularly simple. If $f$ is a monotone Boolean function on $\{0,1\}^{E}$, where $E=\{1,2, \ldots, n\}$, then $c(f)$ is a polygon if and only if $f$ is given by

$$
f\left(x_{1}, x_{2}, \ldots, x_{n}\right)=x_{1} x_{2} \ldots x_{n}
$$

and $c(f)$ is a flower if and only if $f$ is given by

$$
f\left(x_{1}, x_{2}, \ldots, x_{n}\right)=x_{1}+x_{2}+\ldots+x_{n}
$$

It should be remarked that the generalization of the substitution decomposition to arbitrary set families does not, as one might hope, provide a decomposition theory for arbitrary Boolean functions.

The substitution decomposition has been studied previously ([1], $[\mathbf{5}],[\mathbf{2 0}],[\mathbf{4}],[\mathbf{6}]$ ), and a uniqueness theorem has been proved ([20], [4]). It is shown in [20] that a certain uniquely-determined procedure can be carried out; that procedure constructs a factorization of a clutter, and it is easy to see that that factorization is our "standard" factorization. As is shown in [10], this theorem can be derived from the present theory; however, it is not equivalent to our Theorem 25.

Bixby [6] has applied the substitution decomposition to matroids in the following way. Let $e$ be a cell of a non-separable matroid $M=(E, \mathscr{F})$. The clutter $M(e)=(E-e, \mathscr{F}(e))$, where $\mathscr{F}(e)=\{F-e: e \in F \in \mathscr{F}\}$, is called a (matroidal) path clutter. (A fundamental result of [16] states that $M(e)$ determines $M$.) Perhaps surprisingly, the substitution decomposition theory for $M(e)$ can be shown to be equivalent to the theory of Section 6 for $M$. In particular, Bixby shows that $M(e)$ is irreducible if and only if $M$ is 3 -connected.

The cligue clutter of a finite simple graph $G$ is the set family $C(G)=$ ( $V(G), \mathscr{F}$ ), where the members of $\mathscr{F}$ are the vertex-sets of maximal complete subgraphs of $G$. Clearly, $C(G)$ determines $G$ (except for the names of the edges). Moreover, $C(G)$ has no isthmuses. The substitution composition for clique clutters is easily seen to yield the following graph
composition, called by Chvátal [9], "graph substitution". Given graphs $G_{1}, G_{2}$ and $v \in V\left(G_{1}\right), G_{1}\left[G_{2} ; v\right]$ is the graph obtained from $G_{1}$ by replacing $v$ by $G_{2}$ and joining every vertex of $G_{2}$ to every neighbour in $G_{1}$ of $v$. Using Theorem 25, and supplying appropriate definitions, it is easy to derive a unique factorization theorem for graphs, based on this composition; the components of the factorization are irreducible graphs, complete graphs, and edgeless graphs.

## References

1. R. L. Ashenhurst, The decomposition of switching functions, in: Proceedings of the international symposium on the theory of switching (Part 1), (Harvard University Press, Cambridge, 1959).
2. C. Berge, Graphs and hypergraphs (North Holland, Amsterdam, 1973).
3. L. J. Billera, Clutter decomposition and monotonic Boolean functions, Annals of the New York Academy of Sciences 175, Article 1 (1970), 41-48.
4. -On the composition and decomposition of clutters, J. Combinatorial Theory 11 (1971), 234-245.
5. Z. W. Birnbaum and J. D. Esary, Modules of coherent binary systems, SIAM J. Applied Math. 13 (1965), 444-462.
6. R. E. Bixby, Composition and decomposition of matroids and related topics, Thesis, Cornell University (1972).
7. R. E. Bixby and W. H. Cunningham, Matroids, graphs, and 3-connectivity, in: Graph theory and related topics (Academic Press, New York, 1979), 91-103.
8. T. H. Brylawski, A decomposition for combinatorial geometries, Trans. Amer. Math. Soc. 171 (1972), 235-282.
9. V. Chvátal, On certain polytopes associated with graphs, J. Combinatorial Theory B 18 (1975), 138-154.
10. W. H. Cunningham, A combinatorial decomposition theory, Thesis, University of Waterloo (1973).
11. J. Edmonds, Matroid partition, Math. of the Decision Sciences, A.M.S. Lectures in Applied Math. 11 (1968), 335-346.
12.     - Matroid intersection, Annals of Discrete Math. 4 (1979), 39-47.
13. J. Edmonds and R. Giles, A min-max relation for submodular functions on graphs, Annals of Discrete Math. 1 (1977), 185-204.
14. J. E. Hopcroft and R. E. Tarjan, Isomorphism of planar graphs, in: Complexity of computer computations (Plenum Press, New York, 1972).
15. -_Dividing a graph into triconnected components, SIAM J. Computing 2 (1973), 135-158.
16. A. Lehman, $A$ solution of the Shannon switching game, SIAM J. Applied Math. 12 (1964), 687-725.
17. S. Maclane, A structural characterization of planar combinatorial graphs, Duke Math. J. 3 (1937), 460-472.
18. G. J. Minty, On the axiomatic foundations of the theories of directed linear graphs, electrical networks, and network-programming, J. Math. Mech. 15 (1966), 485-520.
19. W. Richardson, Connectivity, decomposition and complexity in matroids, Thesis, University of Waterloo (1973).
20. L. S. Shapley, On committees, in: New methods of thought and procedure (Springer, New York, 1968).
21. C. A. B. Smith, Patroids, J. Combinatorial Theory B 16 (1974), 64-76.
22. B. A. Trakhtenbrot, Towards a theory of non-repeating contact schemes (Russian), Trudi MIAN SSSR 51 (1958), 226-269.
23. W. T. Tutte, Connectivity in graphs (University of Toronto Press, Toronto, 1966).
24.     - Connectivity in matroids, Can. J. Math. 18 (1966), 1301-1324.
25. T. R. S. Walsh, Counting labelled three-connected and homeomorphically irreducible 2-connected graphs, submitted for publication.
26.     - Counting unlabelled three-connected graphs, submitted for publication.
27. D. J. A. Welsh, Matroid theory (Academic Press, London, 1976).
28. H. Whitney, Non-separable and planar graphs, Trans. Amer. Math. Soc. 34 (1932), 339-362.

Carleton University, Ottawa, Ontario;
University of Waterloo, Waterloo, Ontario

