COMBINATORICA 3 (1) (1983) 53—68

DECOMPOSITION OF SUBMODULAR FUNCTIONS

William H. CUNNINGHAM*

Received 27 January 1982

Revised 23 July 1982

A decomposition theory for submodular functions is described. Any such function is shown to have a unique decomposition consisting of indecomposable functions and certain highly decomposable functions, and the latter are completely characterized. Applications include decompositions of hypergraphs based on edge and vertex connectivity, the decomposition of matroids based on three-connectivity, the Gomory—Hu decomposition of flow networks, and Fujishige's decomposition of symmetric submodular functions. Efficient decomposition algorithms are also discussed.

1. Introduction

Let f be a real-valued submodular function defined on subsets of a finite set E. That is, for all $A, B \subseteq E$ we have

$$f(A)+f(B) \geqq f(A \cup B)+f(A \cap B).$$

We also assume, mainly for convenience, that f is *normalized*, that is, $f(\emptyset)=0$. Several examples will be used throughout for illustration, so we introduce them now. (We encourage the reader to check submodularity in each case.)

Example 1. *(Network functions.)* Let G be an undirected graph having vertex set E and suppose we have a positive weight w_j for each edge j. For $A \subseteq E$, let $f(A)$ be $\Sigma(w_j: j$ has at least one end in $A)$. (As a variant of this, we can let $f(A)$ be $\Sigma(w_j: j$ has exactly one end in $A)$.)

Example 2. *(Graph functions.)* Let G be an undirected graph having no isolated vertices and let E be its edge-set. Let $f(A)$ be the cardinality of the set of vertices incident with at least one element of A. (As a generalization of this, we can allow G to be a hypergraph, and/or we can suppose that each vertex has a positive weight with "weight" replacing "cardinality" in the definition of $f(A)$.)

---

* Supported by Songerforschungsbereich 21 DFG, Institut für Operations Research Universität Bonn and by an N.S.E.R.C. of Canada operating grant.

AMS subject classification (1980): 05 C 99; 05 B 35, 68 E 99.

Example 3. *(Matroid functions.)* Let M be a matroid on E and let f be its rank function; that is, for each $A \subseteqq E, f(A)$ is the cardinality of any maximal independent subset of A.

This paper presents a decomposition theory for submodular functions. In the remainder of this section we introduce some basic ideas and handle the simple theory of "direct sum" decomposition. Section 2 contains additional terminology from decomposition theory and the statement of the main results. These assert the uniqueness of a decomposition of any function into indecomposable functions and certain highly decomposable functions, and completely characterize the latter. In Section 3 we derive from these main results unique decomposition theorems for the structures of Examples 1—3; some of these results are new. Section 4 relates the present theory to that of symmetric submodular functions due to Fujishige [8], which was the original motivation for this work. Section 5 contains the proof of the main theorems. The first step is an easy application of the general decomposition theory of Cunningham and Edmonds [5], while the characterization of the highly decomposable functions is somewhat harder to prove. The last section discusses the existence of efficient decomposition algorithms.

We define the *connectivity function* c_f of the function f by

$$c_f(A) = f(A) + f(\bar{A}) - f(E)$$

for all $A \subseteqq E$. (Throughout this paper $\bar{A}$ denotes $E \setminus A$, for any $A \subseteqq E$. For $a, b \in E$ and $A \subseteqq E$, we tend to denote $A \cup \{a\}, A \setminus \{b\}, (A \cup \{a\}) \setminus \{b\}$ by the less cumbersome expressions $A+a, A-b, A+a-b$. We also abbreviate $f(\{a\})$ to $f(a)$.) We have the following easy result.

Proposition 4. *If f is submodular and normalized, then c_f is normalized, submodular, non-negative and symmetric $(c_f(A) = c_f(\bar{A}))$.*

Proof. Clearly c_f is normalized and symmetric. For any $A \subseteqq E$,

$$c_f(A) = f(A) + f(\bar{A}) - f(E) \geqq f(E) + f(\emptyset) - f(E) = 0,$$

so c_f is non-negative. Finally, for $A, B \subseteqq E$

$$\begin{aligned} c_f(A) + c_f(B) &= f(A) + f(\bar{A}) - f(E) + f(B) + f(\bar{B}) - f(E) \\ &\geqq f(A \cap B) + f(A \cup B) + f(\bar{A} \cap \bar{B}) + f(\bar{A} \cup \bar{B}) - 2f(E) \\ &= c_f(A \cap B) + c_f(A \cup B), \end{aligned}$$

so c_f is submodular. ∎

We define $\bar{c}_f$ to be $\min(c_f(A) \colon \emptyset \subset A \subset E)$; $\bar{c}_f = \infty$ if $|E| \leqq 1$. Of course $\bar{c}_f \geqq 0$ and it is appropriate to consider the case when $\bar{c}_f = 0$. (In what follows, we will often drop the subscript from $c, \bar{c}$.) In Examples 1 and 2, this corresponds to the graph G (or hypergraph) being disconnected, and in Example 3 to the matroid M being separable. Of course, these situations lead to natural decompositions of the network, graph or matroid into smaller structures. Before going on to a more substantial notion of decomposition, we dispose of this simpler notion.

Proposition 5. *If* $c(E_1)=0$ *for some non-trivial partition* $\{E_1, E_2\}$ *of* E, *then* f *is the "direct sum" of its restrictions to* E_1 *and* E_2. *That is,* $f(A)=f(A\cap E_1)+f(A\cap E_2)$ *for any* $A\subseteq E$.

Proof. By submodularity, $f(A)\leqq f(A\cap E_1)+f(A\cap E_2)$. But $f(A)-f(A\cap E_1)-f(A\cap E_2)\geqq f(E_1\cup A)-f(E_1)-f(A\cap E_2)\geqq f(E)-f(E_1)-f(E_2)=0$. ∎

By a slight abuse of usual terminology, we call E the *domain* of f. A function f is *non-null* if its domain is different from $\emptyset$, and is *separable* if $\bar{c}=0$. Therefore, a function is separable precisely if it can be expressed as the direct sum of non-null functions. A *separator* of f is a set $A\subseteq E$ having $c(A)=0$. Using Proposition 4, it is easy to show that the set of separators of f is closed under intersection, union, and complementation. Therefore, the minimal non-empty separators of f partition E. Thus we have the following result.

Theorem 6. *Each non-null, normalized, submodular function* f *has a unique expression as the direct sum of non-null, non-separable functions.* ∎

2. Unique decomposition theorems

In view of the elementary theory of direct sum decomposition, we restrict attention to non-separable functions f. Where f is non-separable we call a partition $\{E_1, E_2\}$ of E a *split* of f if $|E_1|\geqq 2\leqq |E_2|$ and $c(E_1)=\bar{c}$. In each of our example functions, a split leads to a decomposition of the underlying structure (and thus the function).

In Example 1, a split corresponds to a minimum weight cut of G which is not induced by a single vertex. In the case when such a cut exists, Gomory and Hu [9] decompose the network into two smaller networks, obtained by shrinking first E_1, then E_2, to a single new vertex e. This is the basic step in their algorithm for finding a minimum cut separating every pair of vertices of a network.

In Example 2, if f is non-separable and $|E|\geqq 2$, then $\bar{c}=1$ or 2, since $c(e)\leqq 2$ for any $e\in E$. If $\bar{c}=2$, G is a non-separable graph, and a split is a *2-separation* of G [13]. That is, $|E_1|\geqq 2\leqq |E_2|$ and there are just two vertices u, v which are incident with elements of both E_1 and E_2. (A non-separable graph has no 2-separation if and only if it is *3-connected* [13].) One can the decompose G into graphs G_1, G_2 where G_i is the subgraph of G induced by E_i plus a single new edge e joining u to v. This decomposition is investigated in [13], [5] and elsewhere. (The case in which $\bar{c}=1$ is not so important; it leads to a variant of the decomposition of graphs into blocks, which can be viewed as a special case of the decomposition of [3].)

In Example 3, if f is non-separable and $|E|\geqq 2$, then $c(e)=1$ for all e, so $\bar{c}=1$, and a split is a 2-separation [14] of the matroid M. (Again, non-separable matroids having no 2-separation are said to be *3-connected* [14].) In this case it is possible to decompose M into matroids M_i on E_i+e, $i=1$ and 2, where $e\notin E$. Every circuit of M is either a circuit of M_i not containing e or is of the form

$(C_1 \cup C_2)-e$ where C_i is a circuit of M_i containing e. The resulting theory of matroid decomposition is described in [5].

Now given a split $\{E_1, E_2\}$ of a non-separable function f on E, we define functions f_1, f_2 on E_1+e, E_2+e respectively, where $e \notin E$, as follows. For any $A \subseteq E_1$, let

$$f_1(A) = f(A),$$
$$f_1(A+e) = f(A \cup E_2) + f(E_1) - f(E),$$

and similarly for f_2. The reader can check that the decompositions of graphs, networks, and matroids suggested above induce precisely this decomposition of the corresponding submodular functions. It should be pointed out that, in contrast to some more familiar notions of decomposition, the functions f_1, f_2 do *not* determine f uniquely. In fact this is not even true for the special classes arising from Examples 1 and 2. We now establish a few of the essential properties of f_1, f_2.

Proposition 7. *For* $i=1$ *and* 2, f_i *is submodular and normalized, and satisfies* $c_i(A)=c(A)$ *for all* $A \subseteq E_i$. *Also* $\bar{c}_i=\bar{c}$ *and, in particular,* f_i *is non-separable.*

Proof. Consider the statement $f_1(A)+f_1(B) \geqq f_1(A \cap B)+f_1(A \cup B)$. If neither A nor B contains e, neither of $A \cap B$ nor $A \cup B$ does, and the result is immediate. Now suppose that one of the sets, say A, contains e and the other does not. Then

$$\begin{aligned} f_1(A)+f_1(B) &= f(A \cup E_2 - e) + f(E_1) - f(E) + f(B) \\ &\geqq f(A \cap B) + f(A \cup B \cup E_2 - e) + f(E_1) - f(E) \\ &= f_1(A \cap B) + f_1(A \cup B). \end{aligned}$$

Finally, if both A and B contain e,

$$\begin{aligned} f_1(A)+f_1(B) &= f(A \cup E_2 - e) + f(B \cup E_2 - e) + 2(f(E_1) - f(E)) \\ &\geqq f((A \cap B) \cup E_2 - e) + f(A \cup B \cup E_2 - e) + 2(f(E_1) - f(E)) \\ &= f_1(A \cap B) + f_1(A \cup B). \end{aligned}$$

Therefore, f_1 is submodular, and similarly so is f_2.

$$\begin{aligned} \text{Now } c_i(A) &= f_1(A) + f_1(E_1 \setminus A + e) - f_1(E_1 + e) \\ &= f(A) + f(\bar{A}) + f(E_1) - f(E_2) - (f(E) + f(E_1) - f(E_2)) \\ &= c(A). \end{aligned}$$

It follows that $\bar{c}_i \geqq \bar{c}$. But $\bar{c}_i = \bar{c}$, because $\bar{c}_i \leqq c_i(E_i) = c(E_i) = \bar{c}$. ∎

Given a split $\{E_1, E_2\}$ of f and f_1, f_2 as above we write $f \to \{f_1, f_2\}$, and call $\{f_1, f_2\}$ a *simple decomposition* of f, corresponding to $\{E_1, E_2\}$ and the *marker* e. A (general) *decomposition* of f is defined inductively to be either $\{f\}$ or a set D' of functions obtained from a decomposition D of f by replacing a member f_1 of D by the members of a simple decomposition of f_1, where the marker of this simple decomposition is not in the domain of any member of D. If D'' is obtained from D by a (non-empty) sequence of operations of the kind described above, then D'' is a *(strict) refinement* of D. If the sequence consists of just one operation, then the refinement is *simple*.

We can associate a graph T with any decomposition D of a function f. The vertices of T are the members of D and the edges of T are the markers of D; each marker joins in T the two members of D whose domains contain it. It is clear that T is a tree. This "decomposition tree" provides a useful way to visualize a decomposition.

Two decompositions D, D' of f are *equivalent* if D' can be obtained from D by replacing some of the markers of D by markers of D'. Notice that this notion of equivalence is stronger than simply requiring that the members of D and D' be "pairwise isomorphic". In particular, it implies that the two decomposition trees are isomorphic. When we make statements about uniqueness of decomposition, we mean uniqueness up to this equivalence. The decomposition D of f is *minimal* with some property P if D has P and there does not exist a decomposition D' of f also having P, such that D is a strict refinement of D'. A decomposition is *trivial* if $|D|=1$. A function f is *prime* if it has no non-trivial decomposition.

The simplest type of unique decomposition result would be one asserting that each non-separable, normalized, submodular function has just one decomposition consisting of prime functions. However, this is not true. Suppose, for example, that f is the matroid function of a rank-one matroid, so $f(A)=1$ for non-empty subsets A of E. Such a function has inequivalent prime decompositions if $|E|$ is at least 4, and has such decompositions having non-isomorphic decomposition trees if $|E|$ is at least 6. In fact, this f has the property that every partition $\{E_1, E_2\}$ of E satisfying $|E_1|\geqq 2\leqq |E_2|$ is a split of f; if $|E|\geqq 4$, we call such a function *brittle*. (Another example of a brittle function is the matroid function of the dual of a rank-one matroid, defined by $f(A)=\min(|A|, |E|-1)$, $A\subseteq E$.) By creating a less symmetrical brittle function, we can exhibit even worse non-uniqueness properties. Suppose that $|E|=4$, fix $B\subseteq E$ with $|B|=2$, and let $f(A)=1+|B\cap A|$, for $\emptyset\neq A\subseteq E$. Then f is brittle, and it is easy to see that it has prime decompositions whose members are not even pairwise isomorphic.

There is another important class of highly decomposable functions, which plays an important role in the decomposition theory. Say that the function f is *semi-brittle* if there is an ordering $e_0, e_1, \dots, e_{n-1}$ of its domain so that the splits of f are precisely the partitions of the form $\{\{e_i, e_{i+1}, \dots, e_{i+j-1}\}, \{e_{i+j}, \dots, e_{i-1}\}\}$, where $0\leqq i\leqq n-1$, $2\leqq j\leqq n-2$, $n\geqq 4$, and subscripts are taken modulo n. As an example of a semi-brittle function, consider the graph function arising from a graph G which is a polygon. (The same function arises as the network function of a polygon in which every edge has weight 1.) By applying the theory of [5], we shall prove the following result.

Theorem 8. *Each normalized, non-separable, submodular function has a unique minimal decomposition, each of whose members is prime, brittle, or semi-brittle.*

It is important to improve Theorem 8 by characterizing the brittle and semi-brittle functions. First, let us observe that it is easy to create more of these functions by combining or perturbing known ones. For example, if f is brittle or semi-brittle, then so is the function αf for any real number $\alpha>0$, where αf is defined by

$(\alpha f)(A)=\alpha f(A)$ for all A. Also, if m is a normalized modular function then $f+m$ is brittle or semi-brittle if f is. (A function m is *modular* if both m and $-m$ are submodular; it is easy to see that every normalized modular function m is determined by its values on singletons, so $m(A)=\sum(m_j: j\in A)$ for all $A\subseteq E$, and numbers $m_j, j\in E$.) Finally, if f_1 and f_2 are brittle or semi-brittle, then so is their sum, provided only that in the case when both are semi-brittle, the semi-brittle orderings must agree. It follows that the brittle and semi-brittle functions are numerous. However, except when $|E|=4$, they all arise via the above operations from three basic functions, as the following result shows.

Theorem 9(a). *Le f be a normalized, non-separable, submodular function with domain E, $|E|\geqq 5$. Then f is brittle or semi-brittle if and only if*

$$f=\alpha_1 f_1+\alpha_2 f_2+\alpha_3 f_3+m$$

where f_1 is the matroid function of a rank-one matroid on E, f_2 is the matroid function of its dual, f_3 is the graph function of a polygon having edge set E, m is a normalized modular function on E, and $\alpha_1, \alpha_2, \alpha_3$ are non-negative reals, not all zero. Moreover, the expression for f is unique, and if f is integer-valued, then $\alpha_1, \alpha_2, \alpha_3$ and m are integer-valued. Finally, f is semi-brittle if and only if $\alpha_3>0$.

Surprisingly, Theorem 9(a) fails for $|E|=4$ (a fact pointed out by É. Tardos), and there exists one more special function. Where $E=\{0, 1, 2, 3\}$, define f_4 by: $f_4(\emptyset)=0$, $f_4(A)=2$ if $|A|=1$, $f_4(A)=3$ if $|A|=2$ and $A\neq\{0, 2\}$, and $f_4(A)=4$ otherwise. It is easy to check that f_4 is non-separable, normalized, submodular and semi-brittle.

Theorem 9(b). *Let f be a normalized, non-separable, submodular function with domain E, $|E|=4$. Then f is brittle or semi-brittle if and only if*

$$f=\alpha_1 f_1+\alpha_2 f_2+\alpha_3 f_3+\alpha_4 f_4+m$$

where $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ are non-negative reals, not all zero, and the semi-brittle orderings for f_3, f_4 agree. Moreover, this expression for f is unique, and if f is integer-valued, then $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ and m are integer-valued. Finally, f is semi-brittle if and only if $\alpha_3+\alpha_4>0$.

3. Applications

In this section we show how unique decomposition theorems for the structures in Examples 1—3 can be derived from Theorems 8 and 9. In each case we have the following general situation. Each member N of a class $\mathcal{N}$ of concrete objects (undirected connected networks, undirected non-separable graphs, non-separable metroids) gives rise to a normalized non-separable submodular function f_N. In order to obtain a unique decomposition result for elements of $\mathcal{N}$, we proceed through the following three steps:

(i) Show that for each $N\in\mathcal{N}, f_N$ determines N;
(ii) Show that for $N\in\mathcal{N}$ if $f_N\to\{f_1, f_2\}$, then there exist $N_1, N_2\in\mathcal{N}$ with $f_i=f_{N_i}$, $i=1$ and 2;
(iii) Use Theorem 9 to determine for which $N\in\mathcal{N}$ f_N is brittle or semi-brittle.

Let us apply this program to Example 3 first. Then (i) is immediate, because the rank function of a matroid determines the matroid. We have already indicated the construction which handles (ii); we omit further details. Now for (iii) suppose that f is a matroid function which is brittle or semi-brittle. For $|E|\geqq 2$ any non-separable matroid function f satisfies $f(E-j)=f(E)$ for all $j\in E$, as does each of f_1, f_2, f_3, so m of Theorem 9 will be zero. For $|E|\geqq 5$ and any $e\in E$ we have $1=f(e)=\alpha_1 f_1(e)+\alpha_2 f_2(e)+\alpha_3 f_3(e)=\alpha_1+\alpha_2+2\alpha_3$. Since f is integer-valued, the α_i are non-negative integers, so the only possibilities are $\alpha_1=1$, $\alpha_2=\alpha_3=0$ and $\alpha_2=1$, $\alpha_1=\alpha_3=0$. Therefore, for $|E|\geqq 5$ there are no semi-brittle matroid functions, and the only brittle ones are f_1 and f_2. It is straightforward using Theorem 9(b) to obtain the same conclusion for $|E|=4$. Let us call the corresponding matroids "bonds" and "polygons". As a consequence of Theorem 8 and the above analysis, we obtain the following unique decomposition result ([5], Theorem 18).

Theorem 10. *Each non-separable matroid M has a unique minimal decomposition each of whose members is 3-connected, a bond or a polygon.* ▮

Now we turn to Example 2, non-separable graphs. In considering (i), we first observe that multiple edges (and loops) are easy to recognize. Next, we observe that a set $A\subseteq E$ is the set of edges incident to a vertex of a simple graph G if and only if A is maximal with the property that $f(A)=|A|+1$ and $f(B)=3$ whenever $B\subseteq A$, $|B|=2$. Therefore, f does determine G uniquely, up to the names of the vertices. We have already described the construction which proves (ii). Finally, we treat (iii). Let f be a brittle or semi-brittle graph function. As before, we can conclude that $m=0$, and for $|E|\geqq 5$ we get $2=f(e)=\alpha_1+\alpha_2+2\alpha_3$, where the α_i must be non-negative integers. So there are four possible solutions, $2f_1$, $2f_2, f_3$, and f_1+f_2. Of course, f_3 is a graph function, arising from a polygon. Also $2f_1$ is a graph function, arising from a bond: a non-separable graph having two vertices. We shall show that the other two are not graph functions. For any two elements $a, b\in E$, $2f_2(\{a, b\})=4$, so if $2f_2$ arises from a graph G, then no two edges of G have a common end; this is impossible if G is non-separable. Similarly, $(f_1+f_2)(\{a, b\})=3$, so if f_1+f_2 arises from a graph G, every two edges have exactly one common end; this is impossible if $|E|\geqq 4$ and G is non-separable. Finally, in the case $|E|=4$ the same reasoning leads to one additional possibility: that f_4 itself is a graph function. But if so, then the graph has 4 vertices, 4 edges, and no triangles, so it has graph function f_3, not f_4. It follows that $2f_1$, f_3 are the only brittle or semi-brittle graph functions. The resulting uniqueness theorem ([5], Theorem 1) is stated below. (This result bears a strong resemblance to Theorem 10, but it is not a consequence of Theorem 10. It is stronger than the specialization of Theorem 10 to "graphic" matroids.)

Theorem 11. *Each non-separable graph has a unique minimal decomposition each of whose members is 3-connected, a bond or a polygon.* ▮

Now we consider Example 1, networks. In order to make a neater unique decomposition theory, we assume that the underlying undirected graphs have no loops or multiple edges. (The former assumption is mainly for convenience.) It is easy to see that a network is determined by its function f, because the weight of the edge (if any) joining $u, v\in E$ is $f(E)-f(E\setminus\{u, v\})$. The construction which demonstrates (ii) has already been described. Now suppose that f is a network function which is brittle or semi-brittle. Since the underlying graph G is loopless, $f(E-j)=f(E)$ for all $j\in E$, so m of Theorem 9 is zero. If $|E|\geqq 5$, we can assume that $f=\alpha_1 f_1+\alpha_2 f_2+\alpha_3 f_3$. (But, unlike the previous cases, we cannot assume that the α_i are integers.) Now any network function satisfies $\sum(f(e)\colon e\in E)=2f(E)$. Therefore, $\sum(\alpha_1+\alpha_2+2\alpha_3\colon e\in E)=2(\alpha_1+\alpha_2(|E|-1)+\alpha_3|E|)$, which implies that $\alpha_1(|E|-2)=\alpha_2(|E|-2)$. Since $|E|\neq 2$, we can conclude that $\alpha_1=\alpha_2$. For $u, v\in E$, let w denote the weight of the edge (if any) joining u to v in the underlying graph. Then $w=f(u)+f(v)-f(\{u, v\})$, so $w=\alpha_1+\alpha_3$ if $f_3(\{u, v\})=3$, and $w=\alpha_1$ otherwise. Now we can easily compute weights of cuts in the underlying graph. The weight of a cut induced by a single vertex is $2\alpha_3+|E|\alpha_1$, and the weight of a cut induced by a pair u, v of vertices satisfying $f_3(\{u, v\})=3$, is $2\alpha_3+(2|E|-2)\alpha_1$. But the latter cannot exceed the former, and $|E|>2$, so $\alpha_1=0$. (A similar argument shows, in the case $|E|=4$, that $\alpha_1=\alpha_2=\alpha_4=0$.) Therefore, $f=\alpha_3 f_3$ and the only brittlre or semi-brittle network functions arise from a graph which is a polygon, each edge of which has the same weight $(=\alpha_3)$. We obtain the following result, which is new, but strongly suggested in the work of Fujishige [8].

Theorem 12. *Each connected, positive-edge-weighted undirected network has a unique minimal decomposition consisting of such networks, in each of which either every minimum capacity cut is induced by a single vertex, or all edges have equal weight and the underlying graph is a polygon.* ∎

The generalization suggested in Example 2, "weighted hypergraphs", actually includes network functions also as a special case. To see this, take the vertices of the hypergraph to be the edges of the network and the edges to be the vertices, keeping the same definition of incidence. (In fact, there is a second equivalent version of the weighted hypergraph function, related to the first by hypergraph duality, in which the roles of edges and vertices are exchanged. The first lends itself to a decomposition based on vertex connectivity and the second to one based on edge connectivity. We continue to consider the first version, but it may be useful to translate some of what follows, including Theorem 13, into the language of the second.) In order that this generalization satisfy (i) we make the assumption that no two vertices u, v of the hypergraph H be incident with the same edges of H. Otherwise, u and v could be replaced by a single vertex having weight equal to the sum of their weights without changing the hypergraph function. For convenience, we also assume that every vertex of H is incident with at least two edges of H. It is true that the function of H determines H, including the vertex weights, up to the names of the vertices. There is a reasonably straightforward inductive proof of this fact, which we omit. The construction which proves (ii) is a natural extension of the one for Example 2. Suppose that $\{E_1, E_2\}$ is a split of H, and $e\notin E$. For $i=1$ and 2 we form a hypergraph H_i having edge-set E_i+e and vertex-set the set of those vertices of H incident with at least one edge in E_i. Each of these vertices has the same weight in H_i as it does in H and each edge in E_i has the same incident vertices in H_i as it does in H. The edge e is incident in H_i to the set of vertices of H incident to at least one edge from each of E_1 and E_2.

Now we want to characterize the weighted hypergraph functions f which are brittle or semi-brittle. By assumption, $f(E-\{j\})=f(E)$ for all $j\in E$, so we can assume that $m=0$ in Theorem 9. Now we can take $E=\{0, 1, 2, \ldots, n-1\}$

where $f_3(\{0, 1\})=f_3(\{1, 2\})=3$. Let S_j be the set of vertices of the hypergraph of f which are incident to j, for each $j\in E$. Where $f=\alpha_1 f_1+\alpha_2 f_2+\alpha_3 f_3$ by Theorem 9, we have $f(\{0, 1, 2\})-f(\{0, 1\})=\alpha_2+\alpha_3=f(\{1, 2\})-f(\{1\})$. Therefore, $S_2\setminus S_1= =S_2\setminus(S_1\cup S_0)$, so $S_2\cap S_0\subseteqq S_1$. It follows from this and the assumption that every vertex is incident with at least two edges that $f(E\setminus\{0, 2\})=f(E)$. That is, $\alpha_1+\alpha_2(n-2)+\alpha_3 n=\alpha_1+\alpha_2(n-1)+\alpha_3 n$, so $\alpha_2=0$. In the case $|E|=4$ the above analysis leads again to the conclusion that $f(\{1, 3\})=f(E)=f(\{0, 2\})$, which implies $\alpha_4=0$, and then $\alpha_2=0$. On the other hand, every function of the form $\alpha_1 f_1+\alpha_3 f_3$ is a weighted hypergraph function. It arises from the hypergraph having vertex-set $\{v, v_0, v_1, \ldots, v_{n-1}\}$ with $S_j=\{v, v_j, v_{j+1}\}$ for each $j\in E=\{0, 1, \ldots, n-1\}$ (with arithmetic modulo n); v has weight α_1 and each v_j has weight α_3. Let us call such a weighted hypergraph a *wheel.* Of course, there are two degenerate wheels. If $\alpha_1=0$, we simply delete v from the vertex-set and from each S_j; if $\alpha_3=0$, then we get a hypergraph having a single vertex v, with each $S_j=\{v\}$. (It is a slightly inconvenient fact that the bond graph does not appear in this generalization. It has been excluded by the assumption that no two vertices should be incident with the same edges, and is replaced by the degenerate hypergraph with $\alpha_1=2, \alpha_3=0$.) As a consequence of Theorem 8 and this analysis we obtain the following result.

Theorem 13. *Each connected weighted hypergraph has a unique minimal decomposition, each of whose members is a prime or a wheel.* ▮

We remark that, just as our results, Theorems 8 and 9, will be derived from the general theory in [5], Theorems 10—13 could also be derived in this way. There are a number of other applications of this theory [3], [5]. It would be interesting to know the extent to which these applications are contained in the theory for submodular functions. Of course, it is possible that there exist applications in which an appropriate submodular function exists in principle, but cannot be defined in a natural way.

4. Symmetric Submodular Functions

A decomposition theory for normalized symmetric submodular functions can be derived from the present theory, by considering the induced decomposition of the connectivity function. Namely, if $f\to\{f_1, f_2\}$, then we write $c_f\to'\{c_{f_1}, c_{f_2}\}$, where $\to'$ is a decomposition relation for symmetric submodular functions. In order to be able to apply this theory to the symmetric variant, it is necessary to show that every normalized symmetric submodular function g is a connectivity function. But this is easy to see: $g=c_f$, where $f=1/2\, g$, for $c(A)=1/2\,(g(A)+g(\bar{A})-g(E))=g(A)$.

The resulting decomposition can be defined directly, as follows. Let g be a normalized symmetric submodular function having domain E. Let $\bar{c}=\min\,(g(A)\colon \emptyset\subset A\subset E)$. If $\bar{c}=0$, then g is *separable* and we obtain it as a direct sum of "smaller" such functions. Otherwise, if $\{E_1, E_2\}$ is a partition of E with $|E_1|\geqq 2\leqq |E_2|$ and $g(E_1)=\bar{c}$, we define a symmetric function g_i with domain $E_i+e, i=1$ and 2, by $g_i(A)=g(A)$, $A\subseteqq E_i$.

This is essentially the approach taken by Fujishige [8], which inspired the present work. This symmetric approach has the disadvantage that it provides weaker uniqueness results: a function is not generally recoverable from its connectivity function. For example, a matroid function and its dual have equal connectivity functions. Therefore, although a decomposition theory for symmetric submodular functions based on this idea may be said to generalize matroid decomposition theory, it also weakens it, since the main unique decomposition theorem for matroids (Theorem 10) is not a corollary.

Applying Proposition 7 and Theorem 8, we obtain the following unique decomposition theorem for symmetric functions. (The reader should be able to supply the missing definitions.)

Theorem 14. *Each normalized, symmetric, non-separable submodular function has a unique minimal decomposition, each of whose members is prime, brittle, or semi-brittle.* ∎

The decomposition of Theorem 14 is closely related to Fujishige's "canonical decomposition" [8]. We can go on to characterize the brittle and semi-brittle functions. For $|E| \geqq 5$, these are the connectivity functions of functions of the form $\alpha_1 f_1 + \alpha_2 f_2 + \alpha_3 f_3 + m$ as in Theorem 9(a). It is easy to see that each such function is a non-negative combination of the connectivity functions of f_1, f_2, f_3, and m. The connectivity function g_1 of f_1 or of f_2 is given by $g_1(\emptyset) = g_1(E) = 0$, $g_1(A) = 1$ for $\emptyset \subset A \subset E$. The connectivity function g_3 of f_3 is given by $g_3(A) =$ number of edges of G having just one end in A, where G is a polygon with vertex-set E. The connectivity function of m is identically zero. Therefore, the resulting characterization is the following.

Theorem 15(a). *Let f be a normalized, non-separable, symmetric submodular function with domain E, $|E| \geqq 5$. Then f is brittle or semi-brittle if and only if $f = \beta_1 g_1 + \beta_3 g_3$ for non-negative reals β_1, β_3 not both zero. Moreover, the expression for f is unique, and f is semi-brittle if and only uf $\beta_3 > 0$.* ∎

For the case $|E| = 4$, we must consider also the connectivity function g_4 of f_4. It is defined by: $g_4(\emptyset) = g_4(E) = 0$, $g_4(\{0, 2\}) = g_4(\{1, 3\}) = 3$, $g_4(A) = 2$ otherwise. We omit the statement of the resulting Theorem 15(b).

5. Proofs

Theorem 8 can be proved quite easily by applying the general theory in Cunningham and Edmonds [5]. Let $\mathscr{F}$ be the class of normalized, non-separable, submodular functions. For each $f \in \mathscr{F}$, $E(f)$ denotes the domain of f and $\to$, as defined in Section 2, is a relation associating elements f of $\mathscr{F}$ with two-element subsets $\{f_1, f_2\}$ of $\mathscr{F}$. A triple such as $(\mathscr{F}, E, \to)$ was defined in [5] to be a *decomposition frame* if four axioms are satisfied. These are F1—F4 of Theorem 16 below; in other words, the content of Theorem 16 is that $(\mathscr{F}, E, \to)$ is a decomposition frame.

Theorem 16. *$(\mathscr{F}, E, \to)$ satisfies:*

F1. *If $f \in \mathscr{F}$ and $f \to \{f_1, f_2\}$, then for some $e \notin E(f)$ and some partition $\{E_1, E_2\}$ of $E(f)$ with $|E_1| \geqq 2 \leqq |E_2|$, we have $E(f_1) = E_1 + e$, $E(f_2) = E_2 + e$.*

F2. *For a split* $\{E_1, E_2\}$ *of* $f\in\mathcal{F}$ *and* $e\notin E(f)$, *there is exactly one simple decomposition* $\{f_1, f_2\}$ *of* f *with marker* e *corresponding to* $\{E_1, E_2\}$. *(In the situation described in* F2 *we denote by* $f(E_i; e)$ *the function* f_i, $i=1$ *and* 2.*)*

F3. *Let* $\{E_1, E_2\}$ *be a split of* $f\in\mathcal{F}$, *let* $A\subset E_1$, *and let* $e\notin E(f)$. *Then* $\{A, E(f)\setminus A\}$ *is a split of* f *if and only if* $\{A, E_1\setminus A+e\}$ *is a split of* $f(E_1; e)$.

F4. *Let* $\{E_1, E_2\}$, $\{E_3, E_4\}$ *be splits of* $f\in\mathcal{F}$ *such that* $E_3\subset E_1$, *and let* $e, g\notin E(f)$, $e\neq g$. *Then* $f(E_1; e)(E_3; g)=f(E_3; g)$, *and* $f(E_1; e)(E_1\setminus E_3+e, g)$ $=f(E_4; g)(E_4\setminus E_2+g; e)$.

Proof. The truth of F1 and F2 is immediate from the definitions, and F3 follows from Proposition 7. To prove F4, let f_i denote $f(E_i; e)$ for $i=1$ and 2, let f_i denote $f(E_i; g)$ for $i=3$ and 4, let f_{13} denote $f_1(E_3; g)$, let f_{14} denote $f_1(E_1\setminus E_3+e; g)$, and let f_{41} denote $f_4(E_4\setminus E_2+g; e)$.

Let $A\subseteq E_3$. Then $f_{13}(A)=f_1(A)=f(A)=f_3(A)$. Moreover,

$$\begin{aligned} f_{13}(A+g) &= f_1((E_1\setminus E_3)\cup A+e)+f_1(E_3)-f_1(E_1+e)\\ &= f(A\cup E_4)+f(E_1)-f(E)+f(E_3)-(f(E)+f(E_1)-f(E))\\ &= f(A\cup E_4)+f(E_3)-f(E)\\ &= f_3(A+g) \end{aligned}$$

Therefore, $f_{13}=f_3$.

Now let $A\subseteq E_1\setminus E_3$. Then $f_{14}(A)=f_1(A)=f(A)=f_4(A)=f_{41}(A)$. Also

$$\begin{aligned} f_{14}(A+g) &= f_1(A\cup E_3)+f_1(E_1\setminus E_3+e)-f_1(E_1+e)\\ &= f(A\cup E_3)+f(E_4)+f(E_1)-f(E)-(f(E)+f(E_1)-f(E))\\ &= f(A\cup E_3)+f(E_4)-f(E)\\ &= f_4(A+g) = f_{41}(A+g). \end{aligned}$$

By symmetry we also have $f_{14}(A+e)=f_{41}(A+e)$.

Now

$$\begin{aligned} f_{14}(A+e+g) &= f_1(A\cup E_3+e)+f_1(E_1\setminus E_3+e)-f_1(E_1+e)\\ &= f(E_2\cup E_3\cup A)+f_1(E_1)-f(E)+f(E_4)+f(E_1)-f(E)\\ &\quad -(f(E)+f(E_1)-f(E))\\ &= f(E_2\cup E_3\cup A)+f(E_1)+f(E_4)-2f(E). \end{aligned}$$

By symmetry the last expression must also be the value of $f_{41}(A+e+g)$, so $f_{41}=f_{14}$, and F4 is proved. ∎

The next result shows that the decomposition frame $(\mathcal{F}, E, \to)$ enjoys an additional important property, called the intersection property ([5]).

Theorem 17. *Let* $\{E_1, E_2\}$ *and* $\{E_3, E_4\}$ *be splits of* $f\in\mathcal{F}$, *such that* $|E_1\cap E_3|\geqq 2$ *and* $E_1\cup E_3\neq E(f)$. *Then* $\{E_1\cap E_3, E_2\cup E_4\}$ *is also a split of* f.

Proof. By Proposition 4, $2\bar{c}=c(E_1)+c(E_3)\geqq(E_1\cap E_3)+c(E_1\cup E_3)\geqq 2\bar{c}$. Thus $c(E_1\cap E_3)=\bar{c}$. Clearly, $|E_1\cap E_3|\geqq 2\leqq|E_2\cup E_4|$, so the result follows. ∎

Theorem 4 of [5] states that any "object" of a decomposition frame having the intersection property has a unique minimal decomposition consisting of prime, brittle, and semi-brittle objects. Thus the present Theorem 8 follows from that result and Theorems 16 and 17. The remainder of the section is devoted to the proof of Theorem 9.

The first step is to find the "modular part" of a brittle or semi-brittle function. Actually, the following result is of more general interest; it indicates that every submodular function has a canonical expression as the sum of a modular function with a "totally normalized polymatroid function". A *polymatroid* function, or β — function [7], is a normalized submodular function which is also increasing: $f(A)\geqq f(B)$ whenever $A\supseteq B$. A function f is *totally normalized* if $f(E-\{j\})=f(E)$ for all $j\in E=E(f)$.

Theorem 18. *Let f be a normalized submodular function with domain E. For each $j\in E$, let $m_j=f(E)-f(E-\{j\})$. Let f' be defined by $f'(A)=f(A)-\sum(m_j\colon j\in A)$ for all $A\subseteq E$. Then f' is a totally normalized polymatroid function. Moreover, the resulting expression $f=f'+m$ is the only expression for f as the sum of a totally normalized polymatroid function with a modular function. Finally, $c_{f'}=c_f$ and if $f\in\mathscr{F}$ and $f\to\{f_1,f_2\}$ then $f'\to\{f_1',f_2'\}$.* ∎

The proof is straightforward, and is omitted. As a concrete example of the process described above consider the variant suggested in Example 1, $f(A)=$ $=\sum(w_j\colon$ exactly one end of j is in $A)$. This f is not increasing, and it is easy to check that f' of Theorem 18 is just $2g$, where g is the network function associated with the same graph.

(An important consequence of Theorem 18 is this: To minimize a general submodular function f, it is enough to minimize a function of the form $f'(A)+\sum(x_j\colon j\in E\setminus A)$, where x is a real-valued vector and f' is a polymatroid function. This observation has also been made by Wolsey [15]. The special case of the latter problem in which f' is a matroid function has been solved by an efficient combinatorial algorithm [4].)

Now we continue the proof of Theorem 9. It is easy to see that a function expressible in the form $f=\alpha_1f_1+\alpha_2f_2+\alpha_3f_3+m$ is brittle or semi-brittle, and is semi-brittle precisely if $\alpha_3>0$. Now suppose that $E=\{0, 1, \ldots, n-1\}$. An *interval* is a set $A\subseteq E$ of the form $\{i, i+1, \ldots, i+j\}$ for $0\leqq i\leqq n-1$, $0\leqq j<n-1$, with arithmetic modulo n. A sequence $A_1, \ldots, A_k$ of disjoint intervals are *consecutive* if $A_i\cup A_{i+1}$ is also an interval, $1\leqq i\leqq k-1$. Clearly, every proper non-empty subset can be expressed uniquely as the union of disjoint, mutually non-consecutive intervals. Since f is brittle or semi-brittle, we may assume that every interval A such that $|A|\geqq 2\leqq|\bar{A}|$ satisfies $c(A)=\bar{c}$. In fact, this must hold for *all* intervals, because for each $j\in E$, $c(j)\leqq c(j-1,j)+c(j,j+1)-c(j-1,j,j+1)=\bar{c}$ by Proposition 4. Moreover, by Theorem 18 we may assume that $f(E-\{j\})=f(E)$ for all $j\in E$, so $f(j)=\bar{c}+f(E)-f(E-\{j\})=\bar{c}$.

We proceed through a sequence of elementary results. The next one can be proved by the technique used in Theorem 17; we omit the details.

Lemma 19. *If* A, B *are intervals such that* $A\cap B\neq\emptyset$ *and* $A\cup B\neq E$, *then* $f(A)+f(B)=f(A\cap B)+f(A\cup B)$. ∎

Lemma 20. *For* $j\in E, f(j, j+1)$ *is independent of* j.

Proof. $f(E)+\bar{c}=f(E-j)+f(j+1)$

$$=f(E-\{j, j-1\})+f(j, j+1) \quad \text{(by Lemma 19)}$$

$$=f(E)+\bar{c}-f(j, j-1)+f(j, j+1).$$

Therefore, $f(j, j+1)=f(j, j-1)$ for all j, and the result follows. ∎

Define p by $f(j, j+1)=\bar{c}+p$; because f is increasing, $p\geqq 0$.

Lemma 21. *For any interval* $A, f(A)=\bar{c}+p(|A|-1)$.

Proof. If $|A|=k+1$, then by Lemma 19, $f(A)=f(B)+f(C)-f(D)$, where B, C, D are intervals with $|B|=|C|=k$, $|D|=k-1$. The result now follows from Lemma 20 by a simple induction. ∎

Lemma 22. *If* A, B, C, D *are consecutive intervals whose union is not* E *such that* $|D|=1$, *then* $f(A\cup C\cup D)=f(A\cup C)+p$.

Proof. $f(A\cup C\cup D)-f(A\cup C)\leqq f(C\cup D)-f(C)=p$, by Lemma 21. But similarly $f(A\cup C\cup D)-f(A\cup C)\geqq f(A\cup B\cup C\cup D)-f(A\cup B\cup C)=p$. ∎

Lemma 23. *If* $|E|\geqq 5$, *then there is a fixed number* q *such that for any* A, $\emptyset\subset A\subset E$, $f(A)=\bar{c}+p(|A|-1)+q(k-1)$, *where* k *is the number of intervals comprising* A. *Moreover,* $0\leqq q\leqq p\leqq \bar{c}-q$.

Proof. For $k=1$, this is just Lemma 21. For $k=2$, it can be proved from Lemma 22, as follows.

Since $|E|\geqq 5$, every set which is comprised of two intervals can be converted into any other by a sequence of single-element additions and deletions as in Lemma 22. By that result, the effect of an addition is to add p to $f(A)$, and the effect of a deletion is to subtract p. It follows in particular that every two-element set A which is not itself an interval has the same f-value; we define q by $f(A)=\bar{c}+p+q$ for such A. Now the result for general A and $k=2$ follows by induction on $|A|$.

We prove the main claim by induction on k. Let $A=\bigcup(A_i\colon 1\leqq i\leqq k)$ where $A_1, B_1, \ldots, A_k, B_k$ is a sequence of consecutive intervals whose union is E. Then $f(A)+f(A_1\cup B_1\cup A_2)\geqq f(A_1\cup A_2)+f(A\cup B_1)$. So

$$f(A) \geqq \bar{c}+p(|A_1|+|A_2|-1)+q+\bar{c}+p(|A|+|B_1|-1)+q(k-2)-$$

$$-(\bar{c}+p(|A_1|+|B_1|+|A_2|-1)) = \bar{c}+p(|A|-1)+q(k-1).$$

Similarly $f(A\cup B_1)+f(A\cup B_2)\geqq f(A)+f(A\cup B_1\cup B_2)$. So

$$f(A) \leqq \bar{c}+p(|A|+|B_1|-1)+q(k-2)+\bar{c}+p(|A|+|B_2|-1)+q(k-2)-$$

$$-(\bar{c}+p(|A|+|B_1|+|B_2|-1)+q(k-3)) = \bar{c}+p(|A|-1)+q(k-1).$$

Finally, we must check the conditions on q. First,

$$\begin{aligned}\bar{c}+p+q &= f(j,j+2)\\ &\geqq f(j,j+1,j+2)+f(j)-f(j,j+1)\\ &= \bar{c}+2p+\bar{c}-(\bar{c}+p)\\ &= \bar{c}+p, \quad \text{so} \quad q\geqq 0.\end{aligned}$$

Since f is increasing, we have $\bar{c}+2p=f(j-1,j,j+1)\geqq f(j-1,j+1)=\bar{c}+p+q$, so $p\geqq q$. Finally, $2\bar{c}=f(j)+f(j+2)\geqq f(j,j+2)+f(\emptyset)=\bar{c}+p+q$, so $p\leqq\bar{c}-q$. ∎

We can now finish the proof of Theorem 9(a). Given a non-empty set $A\subset E$ comprised of k intervals, we have $f_1(A)=1$, $f_2(A)=|A|$ and $f_3(A)=|A|+k$. So choosing $\alpha_3=q$, $\alpha_2=p-q$, $\alpha_1=\bar{c}-p-q$ gives $f(A)=\alpha_1 f_1(A)+\alpha_2 f_2(A)+\alpha_3 f_3(A)$, with $\alpha_1,\alpha_2,\alpha_3\geqq 0$ by Lemma 23, and it is easy to see that this works also for $A=\emptyset$ and E. It is straightforward to check that $\alpha_1,\alpha_2,\alpha_3$ are uniquely determined by considering $A=\{j\}$, $\{j,j+1\}$, $\{j,j+2\}$ and solving. Finally, if f is integer-valued, then clearly so are (m of Theorem 18 and) $\bar{c},p,q$ and hence $\alpha_1,\alpha_2,\alpha_3$. This completes the proof of Theorem 9(a). ∎

The reader who feels by now that the case $|E|=4$ is a nuisance, has my sympathy. To prove Theorem 9(b), we first observe that the stated condition for f to be brittle or semi-brittle is sufficient. Next, we point out that the proof of necessity for $|E|\geqq 5$, up to but not including Lemma 23, applies also to $|E|=4$. Now, where $E=\{0,1,2,3\}$, we may assume that $f(\{0,2\})\geqq f(\{1,3\})$. Let $q=f(\{1,3\})-(\bar{c}+p)$ and let $r=f(\{0,2\})-f(\{1,3\})$. Arguments almost identical to those at the end of the proof of Lemma 23 will show that $q\geqq 0$, $q+r\leqq p$, $p+q+r\leqq\bar{c}$, and (of course) $r\geqq 0$. Choosing $\alpha_1=\bar{c}-p-q-r$, $\alpha_2=p-q-r$, $\alpha_3=q$, $\alpha_4=r$ gives f of the required form, and the α_i are integer-valued if f is. Finally, the uniqueness of the α_i is established by considering the system of equations arising from choosing $A=\{0\}$, $\{0,1\}$, $\{1,3\}$, and $\{0,2\}$.

6. Decomposition Algorithms

In this section we treat briefly the question of the efficient construction of the decompositions described in this paper. We consider both the general case, in which the submodular function f is assumed to be available only via an oracle which can produce $f(A)$ as output, given input A, and special cases like Examples 1 and 2, and concrete versions of Example 3, in which some structure defining f is also available. As usual an algorithm is considered to be efficient if the number of steps is bounded by a polynomial in the problem size. In the general case, the problem size is $|E|$ plus the maximal encoding length of a value of f, and f-evaluations are assumed to be single steps. In the special cases, the problem size is the length of the encoding of the structure defining f.

By Theorem 18, we can assume that f is a polymatroid function. The problem of finding the minimal non-empty separators of a polymatroid function can be solved by a simple variant of the method indicated below for determining a split of a non-separable function. But there is a simpler and much more efficient combinatorial algorithm for this problem [2]. It requires only $O(|E|^2)$ elementary steps, counting f-evaluations as single steps. (Another algorithm having these properties has been suggested by Lovász [11].)

Let us suppose that f is a non-separable polymatroid function and we wish to find a split $\{E_1, E_2\}$ of f, or determine that none exists. An idea introduced in [6] for the matroid case can be used here. Let B, C be disjoint subsets of E and let $E'=E\setminus(B\cup C)$. Define functions f_1, f_2 with domain E' by: $f_1(A)=f(A\cup B)-f(B)$, $f_2(A)=f(A\cup C)-f(C)$, for any $A\subseteq E'$. It is easy to show that f_1, f_2 are polymatroid functions on E'. Notice that $\min\left(f(A)+f(\bar{A})\colon B\subseteq A\subseteq \bar{C}\right)=\min\left(f_1(A')+f_2(E'\setminus A')\colon A'\subseteq E'\right)+f(B)+f(C)$. The latter problem is itself a problem of minimizing a submodular function, and can be solved in polynomial time by a method of Grötschel, Lovász and Schrijver [10], based on the ellipsoid method. (It can also be solved as a "polymatroid intersection problem", but currently the only polynomial-time method for this class of problems also uses the ellipsoid method.) By choosing $|B|=|C|=1$ sufficiently many times and applying the above reduction each time, we can compute $\bar{c}$. Then, choosing $|B|=|C|=2$ sufficiently many times we can compute $\min\left(f(A)+f(\bar{A})\colon A\subseteq E,\ |A|\geqq 2\leqq|\bar{A}|\right)$ and find a split, if one exists.

Now we consider the special functions of Examples 1—3. In each of these cases testing separability is quite easy, and we go on to the problem of finding a split. In Example 1, it is quite easy to use techniques for finding minimum capacity cuts, with slight modifications, to obtain a polynomial-time combinatorial algorithm. In Example 3, there are two known algorithms, [6] and [1], for finding a split. The approach in [6] is essentially the one described above for the general case, except that an ordinary matroid intersection algorithm can be used as a subroutine to obtain an efficient combinatorial algorithm. The algorithm of [1] is a quite different recursive method, and is also more efficient. Tan [12] has further developed this approach. In Example 2 the initial function can be handled by simple graph connectivity methods. To solve the weighted hypergraph problem, we proceed as follows. Construct the bipartite graph corresponding to the hypergraph H and give each vertex $v\in V$ capacity w_v, each vertex $e\in E$ infinite capacity, and each edge infinite capacity. For disjoint sets B, C of E, a partition $\{E_1, E_2\}$ of E with $B\subseteq E_1$, $C\subseteq E_2$ and $f(E_1)+f(E_2)$ minimum can be found by a minimum capacity (vertex) cut calculation. Therefore we can compute $\bar{c}$, and find a split, if one exists, with polynomially many such calculations.

The existence of efficient algorithms to construct a prime decomposition of a given function follows easily from the existence of such algorithms for finding splits. Constructing efficiently the "standard" decomposition of Theorem 8 is harder. One needs to be able to recognize when a decomposition consisting of prime brittle and semi-brittle functions is a strict refinement of another such decomposition. This can be done with the help of Theorem 9. The standard decomposition of f is useful, partly because it provides a representation for *all* the splits of f, even though these may be exponentially many. For example, there exist polynomial-time algorithms to solve the following problems. Given an integer k, find (if possible) a split $\{E_1, E_2\}$ of f such that $|E_1|\geqq k\leqq|E_2|$. Determine the number of splits of f.

Acknowledgement. I am grateful to Professor S. Fujishige, for providing a pre-publication copy of [8] and for suggestions which led to improvements in some of the proofs. I also express my deep appreciation to Dr. Éva Tardos, who discovered an error in the original Theorem 9, leading to the formulation of Theorem 9(b).

References

[1] R. E. BIXBY and W. H. CUNNINGHAM, Matroids, graphs, and 3-connectivity, *in:* Bondy and Murty (eds.) *Graph Theory and Related Topics,* Academic Press, New York (1979), 91—103.

[2] R. E. BIXBY, W. H. CUNNINGHAM and D. M. TOPKIS, The poset of a polymatroid vertex, extreme point, *Report No.* 82250—*OR*, Universität Bonn, 1982.

[3] W. H. CUNNINGHAM, Decomposition of directed graphs, *SIAM J. Algebraic and Discrete Methods* **3** (1982), 214—228.

[4] W. H. CUNNINGHAM, Testing membership in matroid polyhedra, *Report No.* 81207—*OR, Universität Bonn,* 1981, to appear in *J. Combinatorial Th. Ser' B.*

[5] W. H. CUNNINGHAM and J. EDMONDS, A combinatorial decomposition theory, *Canad. J. Math.* **32** (1980), 734—765.

[6] W. H. CUNNINGHAM and J. EDMONDS, Decomposition of linear systems, *in preparation.*

[7] J. EDMONDS, Submodular functions, matroids, and certain polyhedra, *in:* R. K. Guy et al. (eds.) *Combinatorial Structures,* Gordon and Breach, New York (1970), 69—87.

[8] S. FUJISHIGE, Canonical decompositions of symmetric submodular systems, *in:* Saito and Nishizeki (eds.) *Graph Theory and Algorithms, Springer Verlag Lecture Notes in C. S.* **108** (1981), 53—64, also Discr. Applied Math. **5** (1983), 175—190.

[9] R. E. GOMORY and T. C. HU, Multi-terminal network flows, SIAM J. Appl. Math. **9** (1961), 551—570.

[10] M. GRÖTSCHEL, A. SCHRIJVER and L. LOVÁSZ, The ellipsoid method and its consequences in combinatorial optimization, *Combinatorica* **1** (1981), 169—197.

[11] L. LOVÁSZ, *private communication,* 1982.

[12] J. M. TAN, *Matroid 3-Connectivity,* Thesis, Carleton University, 1981.

[13] W. T. TUTTE, *Connectivity in Graphs,* University of Toronto Press, 1966.

[14] W. T. TUTTE, Connectivity in matroids, *Canad. J. Math.* **18** (1966), 1301—1324.

[15] L. A. WOLSEY, *private communication,* 1981.

William H. Cunningham

Department of Mathematics and Statistics
Carleton University, Ottawa, Canada