Part IV

Matroids and Submodular Functions

Part IV: Matroids and Submodular Functions

Matroids form an important tool in combinatorial optimization. Among other, they apply to shortest and disjoint trees in undirected graphs, to bipartite matching, and to directed cut covering.

Matroids were introduced by Whitney in 1935, and equivalent axiom systems were considered in the 1930s by Nakasawa, Birkhoff, and van der Waerden. They were motivated by questions from algebra, geometry, and graph theory. The importance of matroids for combinatorial optimization was revealed by J. Edmonds in the 1960s, who found efficient algorithms and min-max relations for optimization problems involving matroids.

Matroids are exactly those structures where the greedy algorithm yields an optimum solution. Edmonds discovered that matroids have an even stronger algorithmic property: also optimization over intersections of two different matroids can be done efficiently. It is closely related to matroid union. Among the consequences of matroid intersection and union methods and results are min-max relations, polyhedral characterizations, and algorithms for bipartite matching, common transversals, and tree packing and covering. (In fact, tree packing and covering are best investigated within the structures offered by matroids. This insight was obtained already in the original paper of Nash-Williams on tree packing. That is why we discuss matroids before Part V on trees and forests.)

While bipartite matching is generalized by matroid intersection, nonbipartite matching is generalized by *matroid matching*. We prove in Chapter 43 Lovász's matroid matching theorem for linear matroids. For general matroids the problem is intractable.

The rank function of a matroid is a special case of a submodular function. Submodular functions give rise to a polyhedral generalization of matroids, the *polymatroids*. Most of matroid theory can be lifted to the level of submodular functions and polymatroids. Next to having applications by its own, it will also be used in Part V where we consider submodular functions defined on digraphs (Chapter 60). This applies to directed variants of tree and cut packing and covering, and to graph orientation and connectivity augmentation.

Chapters:

Chapter 39

Matroids

This chapter gives the basic definitions, examples, and properties of matroids. We use the shorthand notation

$$X + y := X \cup \{y\} \text{ and } X - y := X \setminus \{y\}.$$

39.1. Matroids

A pair $(S, \mathcal{I})$ is called a *matroid* if S is a finite set and $\mathcal{I}$ is a nonempty collection of subsets of S satisfying:

(39.1) (i) if $I \in \mathcal{I}$ and $J \subseteq I$, then $J \in \mathcal{I}$,
(ii) if $I, J \in \mathcal{I}$ and $|I| < |J|$, then $I + z \in \mathcal{I}$ for some $z \in J \setminus I$.

(These axioms are given by Whitney [1935].)

Given a matroid $M = (S, \mathcal{I})$, a subset I of S is called *independent* if I belongs to $\mathcal{I}$, and *dependent* otherwise. For $U \subseteq S$, a subset B of U is called a *base* of U if B is an inclusionwise maximal independent subset of U. That is, $B \in \mathcal{I}$ and there is no $Z \in \mathcal{I}$ with $B \subset Z \subseteq U$.

It is not difficult to see that, under condition (39.1)(i), condition (39.1)(ii) is equivalent to:

(39.2) for any subset U of S, any two bases of U have the same size.

The common size of the bases of a subset U of S is called the *rank* of U, denoted by $r_M(U)$. If the matroid is clear from the context, we write $r(U)$ for $r_M(U)$.

A set is called simply a *base* if it is a base of S. The common size of all bases is called the *rank* of the matroid. A subset of S is called *spanning* if it contains a base as a subset. So bases are just the inclusionwise minimal spanning sets, and also just the independent spanning sets. A *circuit* of a matroid is an inclusionwise minimal dependent set. A *loop* is an element s such that $\{s\}$ is a circuit. Two elements s, t of S are called *parallel* if $\{s, t\}$ is a circuit.

Nakasawa [1935] showed the equivalence of axiom system (39.1) with an ostensibly weaker system, which will be useful in proofs:

Theorem 39.1. *Let S be a finite set and let $\mathcal{I}$ be a nonempty collection of subsets satisfying* (39.1)(i)*. Then* (39.1)(ii) *is equivalent to:*

(39.3) *if $I, J \in \mathcal{I}$ and $|I \setminus J| = 1$, $|J \setminus I| = 2$, then $I + z \in \mathcal{I}$ for some $z \in J \setminus I$.*

Proof. Obviously, (39.1)(ii) implies (39.3). Conversely, (39.1)(ii) follows from (39.3) by induction on $|I \setminus J|$, the case $|I \setminus J| = 0$ being trivial. If $|I \setminus J| \geq 1$, choose $i \in I \setminus J$. We apply the induction hypothesis twice: first to $I - i$ and J to find $j \in J \setminus I$ with $I - i + j \in \mathcal{I}$, and then to $I - i + j$ and J to find $j' \in J \setminus (I + j)$ with $I - i + j + j' \in \mathcal{I}$. Then by (39.3) applied to I and $I - i + j + j'$, we have that $I + j \in \mathcal{I}$ or $I + j' \in \mathcal{I}$. ∎

39.2. The dual matroid

With each matroid M, a dual matroid M^* can be associated, in such a way that $(M^*)^* = M$. Let $M = (S, \mathcal{I})$ be a matroid, and define

(39.4) $\mathcal{I}^* := \{I \subseteq S \mid S \setminus I \text{ is a spanning set of } M\}$.

Then (Whitney [1935]):

Theorem 39.2. *$M^* = (S, \mathcal{I}^*)$ is a matroid.*

Proof. Condition (39.1)(i) trivially holds for $\mathcal{I}^*$. To see (39.1)(ii), consider $I, J \in \mathcal{I}^*$ with $|I| < |J|$. By definition of $\mathcal{I}^*$, $S \setminus J$ contains some base B of M. As also $S \setminus I$ contains some base of M, and as $B \setminus I \subseteq S \setminus I$, there exists a base B' of M with $B \setminus I \subseteq B' \subseteq S \setminus I$. Then $J \setminus I \not\subseteq B'$, since otherwise (as $B \cap I \subseteq I \setminus J$, and as $B \setminus I$ and $J \setminus I$ are disjoint, since $B \cap J = \emptyset$)

(39.5) $|B| = |B \cap I| + |B \setminus I| \leq |I \setminus J| + |B \setminus I| < |J \setminus I| + |B \setminus I| \leq |B'|$,

which is a contradiction. As $J \setminus I \not\subseteq B'$, there is a $z \in J \setminus I$ with $z \notin B'$. So B' is disjoint from $I + z$. Hence $I + z \in \mathcal{I}^*$. ∎

The matroid M^* is called the *dual matroid* of M. The bases of M^* are precisely the complements of the bases of M. This implies $(M^*)^* = M$, which justifies the name dual.

Theorem 39.3. *The rank function r_{M^*} of the dual matroid M^* satisfies, for $U \subseteq S$:*

(39.6) $r_{M^*}(U) = |U| + r_M(S \setminus U) - r_M(S)$.

Proof. Let $\mathcal{B}$ and $\mathcal{B}^*$ denote the collections of bases of M and of M^*, respectively. Then

(39.7) $$r_{M^*}(U) = \max\{|U \cap A| \mid A \in \mathcal{B}^*\} = \max\{|U \setminus B| \mid B \in \mathcal{B}\}$$
$$= |U| - \min\{|B \cap U| \mid B \in \mathcal{B}\}$$
$$= |U| - r_M(S) + \max\{|B \setminus U| \mid B \in \mathcal{B}\}$$
$$= |U| - r_M(S) + r_M(S \setminus U).$$ ∎

The circuits of M^* are called the *cocircuits* of M. They are the inclusionwise minimal sets intersecting each base of M (as they are the inclusionwise minimal sets contained in no base of M^*, that is, not contained in the complement of any base of M). The loops of M^* are the *coloops* or *bridges* of M, and parallel elements of M^* are called *coparallel* or *in series* in M.

Let $M = (S, \mathcal{I})$ be a matroid, and suppose that we can test in polynomial time if any subset of S is independent in M (or we have an oracle for that). Then we can calculate, for any subset U of S, the rank $r_M(U)$ of U in polynomial time (by growing an independent set (starting from $\emptyset$) to an inclusionwise maximal independent subset of U). It follows that we can test in polynomial time if any subset U of S in independent in M^*, just by testing if $r_M(S \setminus U) = r_M(S)$.

A matroid $M = (S, \mathcal{I})$ is called *connected* if $r_M(U) + r_M(S \setminus U) > r_M(S)$ for each nonempty proper subset U of S. This is equivalent to: for any two elements $s, t \in S$ there exists a circuit containing both s and t. One may derive from (39.6) that a matroid M is connected if and only if M^* is connected.

39.3. Deletion, contraction, and truncation

We can derive matroids from matroids by 'deletion' and 'contraction'. Let $M = (S, \mathcal{I})$ be a matroid and let $Y \subseteq S$. Define

(39.8) $$\mathcal{I}' := \{Z \mid Z \subseteq Y, Z \in \mathcal{I}\}.$$

Then $M' = (Y, \mathcal{I}')$ is a matroid again, as directly follows from the matroid axioms (39.1). M' is called the *restriction* of M to Y, denoted by $M|Y$. If $Y = S \setminus Z$ with $Z \subseteq S$, we say that M' arises by *deleting* Z, and denote M' by $M \setminus Z$. Clearly, the rank function of $M|Y$ is the restriction of the rank function of M to subsets of Y.

Contraction is the operation dual to deletion. *Contracting* Z means replacing M by $(M^* \setminus Z)^*$. This matroid is denoted by M/Z. If $Y = S \setminus Z$, then we denote $M \cdot Y := M/Z$. Theorem 39.3 implies that the rank function r' of M/Z satisfies

(39.9) $$r_{M/Z}(X) = r(X \cup Z) - r(Z)$$

for $X \subseteq S \setminus Z$.

We can describe contraction as follows. Let $Z \subseteq S$ and let X be a base of Z. Then

(39.10) a subset I of $S \setminus Z$ is independent in M/Z if and only if $I \cup X$ is independent in M.

Note that for disjoint subsets Y, Z of S one has $(M \setminus Y) \setminus Z = M \setminus (Y \cup Z)$ and hence $(M/Y)/Z = M/(Y \cup Z)$. Moreover, deletion and contraction commute, as for any two distinct $x, y \in S$ and any $Z \subseteq S \setminus \{x, y\}$ one has (using (39.9)):

(39.11) $$r_{M\setminus x/y}(Z) = r_{M\setminus x}(Z \cup \{y\}) - r_{M\setminus x}(\{y\}) = r_M(Z \cup \{y\}) - r_M(\{y\}) = r_{M/y}(Z) = r_{M/y\setminus x}(Z).$$

If matroid M' arises from M by a series of deletions and contractions, M' is called a *minor* of M.

The circuits of $M|Y$ are exactly the circuits of M contained in Y, and the circuits of $M \cdot Y$ are exactly the minimal nonempty sets $C \cap Y$, where C is a circuit of M.

Another operation is that of 'truncation'. Let $M = (S, \mathcal{I})$ be a matroid and let k be a natural number. Define $\mathcal{I}' := \{I \in \mathcal{I} \mid |I| \leq k\}$. Then $(S, \mathcal{I}')$ is again a matroid, called the *k-truncation* of M.

39.4. Examples of matroids

We describe some basic classes of matroids.

Uniform matroids. An easy class of matroids is given by the *uniform matroids*. They are determined by a set S and a number k: the independent sets are the subsets I of S with $|I| \leq k$. This trivially gives a matroid, called a *k-uniform matroid* and denoted by U_n^k, where $n := |S|$.

Linear matroids (Grassmann [1862], Steinitz [1913]). Let A be an $m \times n$ matrix. Let $S := \{1, \ldots, n\}$ and let $\mathcal{I}$ be the collection of all those subsets I of S such that the columns of A with index in I are linearly independent. That is, such that the submatrix of A consisting of the columns with index in I has rank $|I|$.

Then $(S, \mathcal{I})$ is a matroid (property (39.1)(ii) was proved by Grassmann [1862] and by Steinitz [1913], and is called *Steinitz' exchange property*). Condition (39.1)(i) is trivial. To see condition (39.1)(ii), let $I, J \in \mathcal{I}$ with $|I| < |J|$. Then I spans an $|I|$-dimensional space $\overline{I}$. So $J \not\subseteq \overline{I}$. Take $j \in J \setminus \overline{I}$. Then $I + j \in \mathcal{I}$ and $j \in J \setminus I$.

Any matroid obtained in this way, or isomorphic to such a matroid, is called a *linear matroid*. If A has entries in a field $\mathbb{F}$, then M is called *representable over* $\mathbb{F}$. We will also say that M is *represented by* (the columns of) A, and A is called a *representation* of M.

Note that the rank $r_M(U)$ of any subset U of S is equal to the rank of the matrix formed by the columns indexed by U.

The dual matroid of a matroid representable over a field $\mathbb{F}$ is again representable over $\mathbb{F}$. Indeed, we can assume that the matrix A is of the form $[I_m \;\; B]$, where I_m is the $m \times m$ identity matrix, and B is an $m \times (n - m)$

matrix. Then the dual matroid can be represented by the matrix $[B^{\mathsf{T}}\ \ I_{n-m}]$, as follows directly from elementary linear algebra. This implies that the class of matroids representable over $\mathbb{F}$ is closed under taking minors.

MacLane [1936] (and also Lazarson [1958]) showed that nonlinear matroids exist.

Binary matroids. A matroid representable over GF(2) — the field with two elements — is called a *binary matroid*. For later purposes, we give some characterizations of binary matroids. The following is direct (Whitney [1935]):

(39.12) a matroid M is binary if and only if for each choice of circuits $C_1, \ldots, C_t$, the set $C_1 \triangle \cdots \triangle C_t$ can be partitioned into circuits.

In a binary matroid M, disjoint unions of circuits are called the *cycles* of M. Of special interest is the *Fano matroid* F_7, represented by the nonzero vectors in $\mathrm{GF}(2)^3$.

Tutte [1958a,1958b] showed that the unique minor-minimal nonbinary matroid is U_4^2, the 2-uniform matroid on 4 elements. (We follow the proof suggested by A.M.H. Gerards.)

Theorem 39.4. *A matroid is binary if and only if it has no U_4^2 minor.*

Proof. Necessity follows from the facts that the class of binary matroids is closed under taking minors and that U_4^2 is not binary.

To see sufficiency, we first show the following. Let M and N be matroids on the same set S. Call a set *wrong* if it is a base of precisely one of M and N. A *far base* is a common base B of M and N such that there is no wrong set X with $|B \triangle X| = 2$. We first show:

(39.13) if M and N are different and have a far base, then M or N has a U_4^2 minor.

Let M, N form a counterexample with S as small as possible. Let B be a far base and X be a wrong set with $|B \triangle X|$ minimal. Then $B \cup X = S$, since we can delete $S \setminus (B \cup X)$. Similarly (by considering M^* and N^*), $B \cap X = \emptyset$. Then, by the minimality of $|B \triangle X|$, X is the only wrong set. By symmetry, we may assume that X is a base of M. Then M has a base B' with $|B \triangle B'| = 2$. By the uniqueness of X, B' is also a base of N. By the minimality of $|B \triangle X|$, B' is not far. Hence, by the uniqueness of X, $|B' \triangle X| = 2$. So $|S| = 4$.

Let $S = \{a, b, c, d\}$, $B = \{a, b\}$, $X = \{c, d\}$. Since $M \neq U_4^2$ by assumption, we may assume that $\{a, c\}$ is not a base of M. Hence, since $\{a\}$ and $\{c, d\}$ are independent in M, $\{a, d\}$ is a base of M. Similarly, since $\{c\}$ and $\{a, b\}$ are independent in M, $\{b, c\}$ is a base of M.

Since B is far, $\{a, d\}$ and $\{b, c\}$ are bases also of N, and $\{a, c\}$ is not a base of N. So $\{c\}$ is independent in N, implying that $\{c, a\}$ or $\{c, d\}$ is a base of N, a contradiction. This proves (39.13).

Now let M be a nonbinary matroid on a set S. Choose a base B of M. Let $\{x_b \mid b \in B\}$ be a collection of linearly independent vectors over GF(2). For each $s \in S \setminus B$, let C_s be the circuit contained in $B \cup \{s\}$, and define

$$(39.14) \qquad x_s := \sum_{b \in C_s \setminus \{s\}} x_b.$$

Let N be the binary matroid represented by $\{x_s \mid s \in S\}$. Now for each $b \in B$ and each $s \in S \setminus B$ one has that $(B \setminus \{b\}) \cup \{s\}$ is a base of M if and only if it is a base of N. So B is a far base. Since N is binary, we know that $N \neq M$ and that N has no U_4^2 minor. Hence, by (39.13), M has a U_4^2 minor. ∎

Regular matroids. A matroid is called *regular* if it is representable over each field. It is equivalent to requiring that it can be represented over $\mathbb{R}$ by the columns of a totally unimodular matrix.

Regular matroids are characterized by Tutte [1958a,1958b] as those binary matroids not having an F_7 or F_7^* minor. (Gerards [1989b] gave a short proof.)

A basic decomposition theorem of Seymour [1980a] states that each regular matroid can be obtained by taking 1-, 2-, and 3-sums from graphic and cographic matroids and from copies of a 10-element matroid called R_{10}. (We do not use this theorem in this book. Background can be found in the book of Truemper [1992].)

Algebraic matroids (Steinitz [1910]). Let L be a field extension of a field K and let S be a finite subset of L. Let $\mathcal{I}$ be the collection of all subsets $\{s_1, \ldots, s_n\}$ of S that consist of algebraically independent elements over K. That is, there is no nonzero polynomial $p(x_1, \ldots, x_n) \in K[x_1, \ldots, x_n]$ with $p(s_1, \ldots, s_n) = 0$. Then $(S, \mathcal{I})$ is a matroid, and matroids arising in this way are called *algebraic* (*over* K). (Steinitz [1910] showed that $(S, \mathcal{I})$ satisfies the matroid axioms, although the term matroid was not yet introduced.)

To see that $(S, \mathcal{I})$ is a matroid, we check (39.3). It suffices to show that for all $s_1, \ldots, s_n \in S$ one has:

(39.15) if $\{s_1, s_2, s_3, \ldots, s_{n-1}\} \in \mathcal{I}$ and $\{s_3, \ldots, s_{n-1}, s_n\} \in \mathcal{I}$, then $\{s_1, s_3, \ldots, s_n\} \in \mathcal{I}$ or $\{s_2, s_3, \ldots, s_n\} \in \mathcal{I}$.

Suppose not. Then there exist nonzero polynomials $p(x_1, x_3, \ldots, x_n)$ and $q(x_2, x_3, \ldots, x_n)$ over K with $p(s_1, s_3, \ldots, s_n) = 0$ and $q(s_2, s_3, \ldots, s_n) = 0$. We may assume that p and q are irreducible. Moreover, since $\{s_3, \ldots, s_n\} \in \mathcal{I}$, p and q are relatively prime. Define $F := K(x_1, x_2, \ldots, x_{n-1})$. So p and q belong to the Euclidean ring $F[x_n]$. Let r be the g.c.d. of p and q in $F[x_n]$. As p and q are relatively prime, we know $r \in F$, and hence we may assume $r \in K[x_1, \ldots, x_{n-1}]$. Now $r = \alpha p + \beta q$ for some $\alpha, \beta \in F[x_n]$. So $r(s_1, \ldots, s_{n-1}) = 0$, contradicting the fact that $\{s_1, \ldots, s_{n-1}\} \in \mathcal{I}$. This proves (39.15).

Each linear matroid is algebraic (as we can consider the linear relations between the elements as polynomials of rank 1), while Ingleton [1971] gave an example of a nonlinear algebraic matroid. Examples of nonalgebraic matroids were given by Ingleton and Main [1975] and Lindström [1984,1986]. The class of algebraic matroids can be easily seen to be closed under taking minors (deletion is direct, while contraction of an element t corresponds to replacing K by $K(t)$), but it is unknown if it is closed under duality.

In fact, for any field K, the class of matroids that are algebraic over K is closed under taking minors, since Lindström [1989] showed that any matroid algebraic over $K(t)$ (for any t), is also algebraic over K.

For an in-depth survey on algebraic matroids, see Oxley [1992].

Graphic matroids (Birkhoff [1935c], Whitney [1935]). Let $G = (V, E)$ be a graph and let $\mathcal{I}$ be the collection of all subsets of E that form a forest. Then $M = (E, \mathcal{I})$ is a matroid. Condition (39.1)(i) is trivial. To see that condition (39.2) holds, let $F \subseteq E$. Then, by definition, each base U of F is an inclusionwise maximal forest contained in F. Hence U forms a spanning tree in each component of the graph (V, F). So U has $|V| - k$ elements, where k is the number of components of (V, F). So each base of F has $|V| - k$ elements, proving (39.2).

The matroid M is called the *cycle matroid* of G, denoted by $M(G)$. Any matroid obtained in this way, or isomorphic to such a matroid, is called a *graphic matroid*.

Trivially, the circuits of $M(G)$, in the matroid sense, are exactly the circuits of G, in the graph sense. The bases of $M(G)$ are exactly the inclusionwise maximal forests F of G. So if G is connected, the bases are the spanning trees.

The rank function of $M(G)$ can be described as follows. For each subset F of E, let $\kappa(V, F)$ denote the number of components of the graph (V, F). Then for each $F \subseteq E$:

(39.16) $$r_{M(G)}(F) = |V| - \kappa(V, F).$$

Note that deletion and contraction in the matroid correspond to deletion and contraction of edges in the graph.

Graphic matroids are regular, that is, representable over any field: orient the edges of G arbitrarily, and consider the $V \times E$ matrix L given by: $L_{v,e} = +1$ if v is the head of e, $L_{v,e} := -1$ if v is the tail of e, and $L_{v,e} := 0$ otherwise (for $v \in V$, $e \in E$). Then a subset F of E is a forest if and only if the set of columns with index in F is linearly independent.

By a theorem of Tutte [1959], the graphic matroids are precisely those regular matroids containing no $M(K_5)^*$ and $M(K_{3,3})^*$ minor. (Alternative proofs were given by Ghouila-Houri [1964] (Chapitre III), Seymour [1980d], Truemper [1985], Wagner [1985], and Gerards [1995b].)

Cographic matroids (Whitney [1935]). The dual of the cycle matroid $M(G)$ of a graph $G = (V, E)$ is called the *cocycle matroid* of G, and denoted by $M^*(G)$. Any matroid obtained in this way, or isomorphic to such a matroid, is called a *cographic matroid*.

So the bases of $M^*(G)$ are the complements of maximal forests of G. (So if G is connected, these are exactly the complements of the spanning trees in G.)

Hence the independent sets are those edge sets F for which $E \setminus F$ contains a maximal forest of G; that is, $(V, E \setminus F)$ has the same number of components as G.

A subset C of E is a circuit of $M^*(G)$ if and only if C is an inclusionwise minimal set with the property that $(V, E \setminus C)$ has more components than G. Hence C is a circuit of $M^*(G)$ if and only if C is an inclusionwise minimal nonempty cut in G.

The rank function of $M^*(G)$ can be described as follows. Again, for each subset F of E, let $\kappa(V, F)$ denote the number of components of the graph (V, F). Then (39.6) and (39.16) give that for each $F \subseteq E$:

(39.17) $$r_{M^*(G)}(F) = |F| - \kappa(V, E \setminus F) + \kappa(V, E).$$

Let G be an (embedded) planar graph, and let G^* be the dual planar graph of G. Then the cycle matroid $M(G^*)$ of G^* is isomorphic to the cocycle matroid $M^*(G)$ of G.

A theorem of Whitney [1933] implies that a matroid is both graphic and cographic if and only if it is isomorphic to the cycle matroid of a planar graph.

Transversal matroids (Edmonds and Fulkerson [1965], Mirsky and Perfect [1967]). Let $\mathcal{X} = (X_1, \ldots, X_n)$ be a family of subsets of a finite set S and let $\mathcal{I}$ be the collection of all partial transversals of $\mathcal{X}$. Then $M = (S, \mathcal{I})$ is a matroid, as follows directly from Corollary 22.4a. Any matroid obtained in this way, or isomorphic to such a matroid, is called a *transversal matroid* (*induced by* $\mathcal{X}$).

The bases of this matroid are the inclusionwise maximal partial transversals. If $\mathcal{X}$ has a transversal, the bases of M are the transversals of $\mathcal{X}$. In fact, Theorem 22.5 implies that we can assume the latter situation:

(39.18) Let M be the transversal matroid induced by the family $\mathcal{X}$. Then $\mathcal{X}$ has a subfamily $\mathcal{Y}$ such that M is equal to the transversal matroid induced by $\mathcal{Y}$ and such that $\mathcal{Y}$ has a transversal.

So we can assume that any transversal matroid has the transversals of a family of sets as bases.

It follows from Kőnig's matching theorem that the rank function r of the transversal matroid induced by $\mathcal{X}$ is given by

(39.19) $$\begin{aligned} r(U) &= \min_{T \subseteq U}(|U \setminus T| + |\{i \mid X_i \cap T \neq \emptyset\}|) \\ &= \min_{I \subseteq \{1,\ldots,n\}}(n - |I| + \big|\bigcup_{i \in I}(X_i \cap U)\big|) \end{aligned}$$

for $U \subseteq S$. This follows directly from Theorem 22.2 and Corollary 22.2a, applied to the family $(X_1 \cap U, \ldots, X_n \cap U)$.

Piff and Welsh [1970] (cf. Atkin [1972]) showed that

(39.20) any transversal matroid is representable over all fields, except for finitely many finite fields.

If the sets $X_1, \ldots, X_m$ form a partition of S, one speaks of a *partition matroid*. Trivially, each partition matroid is graphic and cographic (by considering a graph consisting of vertex-disjoint parallel classes of edges). Also uniform matroids are special cases of transversal matroids.

Gammoids (Perfect [1968]). An extension of transversal matroids is obtained by taking a directed graph $D = (V, A)$ and subsets U and S of V. For $X, Y \subseteq V$, call X *linked to* Y if $|X| = |Y|$ and D has $|X|$ vertex-disjoint $X - Y$ paths. (So X is the set of starting vertices of these paths, and Y the set of end vertices.)

Let $\mathcal{I}$ be the collection of subsets I of S such that some subset of U is linked to I. Then $M = (S, \mathcal{I})$ is a matroid. This follows from Theorem 9.11: let $I, J \in \mathcal{I}$ with $|I| < |J|$. Let $T := I \cup J$. Let k be the maximum number of disjoint $U - T$ paths. So $k \geq |J| > |I|$. By Theorem 9.11, there exist k disjoint $U - T$ paths covering I. Hence $I + j \in \mathcal{I}$ for some $j \in J \setminus I$. So M is a matroid.

Matroids obtained in this way are called *gammoids*. If $S = V$, the gammoid is called a *strict gammoid* (*induced by* D, U). Hence:

(39.21) gammoids are exactly the restrictions of strict gammoids.

The bases of the strict gammoid induced by D, U are the subsets B of V such that U is linked to B. In particular, U is a base.

From Menger's theorem (Corollary 9.1a) one easily derives the following formula for the rank function r_M of M:

(39.22) $$r_M(X) = \min\{|Y| \mid Y \text{ intersects each } U - X \text{ path}\}$$

for $X \subseteq S$. (One may prove easily that the right-hand side of (39.22) satisfies Theorem 39.8 below, thus proving again that M is a matroid.)

39.4a. Relations between transversal matroids and gammoids

Ingleton and Piff [1973] showed the following theorem (based on a duality of bipartite graphs and directed graphs similar to that described in Section 16.7c). The proof provides an alternative proof that gammoids are indeed matroids.

Theorem 39.5. *Strict gammoids are exactly the duals of the transversal matroids.*

Proof. Let M be the strict gammoid induced by the directed graph $D = (V, A)$ and $U \subseteq V$. We can assume that $(v, v) \in A$ for each $v \in V$. For each $v \in V$, let

(39.23) $$X_v := \{u \in V \mid (u, v) \in A\}.$$

Let L be the transversal matroid induced by the family $\mathcal{X} := (X_v \mid v \in V \setminus U)$. We show that $L = M^*$.

As $v \in X_v$ for each $v \in V \setminus U$, the set $V \setminus U$ is a transversal of $\mathcal{X}$. Hence the bases of L are the transversals of $\mathcal{X}$. As U is a base of the strict gammoid induced by D, U, it suffices to show, for each $B \subseteq V$:

(39.24) U is linked to B in D if and only if $V \setminus B$ is a transversal of $\mathcal{X}$.

To see necessity in (39.24), let U be linked to B in D and let $\mathcal{P}$ be a set of $|U|$ disjoint $U - B$ paths. Then for each $v \in V \setminus U$, let $x_v := u$ if v is entered by an arc (u, v) in a path P in $\mathcal{P}$ and let $x_v := v$ otherwise. Then:

(39.25) (i) $x_v \in X_v$, (ii) $x_v \neq x_{v'}$ for $v \neq v' \in V \setminus U$, and (iii) $\{x_v \mid v \in V \setminus U\} = V \setminus B$.

So $V \setminus B$ is a transversal of $\mathcal{X}$.

To see sufficiency in (39.24), let $V \setminus B$ be a transversal of $\mathcal{X}$. Hence there exist x_v for $v \in V \setminus U$ satisfying (39.25). Let A' be the set of arcs (x_v, v) of D with $v \in V \setminus U$. Then $V \setminus U$ is the set of vertices entered by an arc in A', and $V \setminus B$ is the set of vertices left by an arc in A'. Hence U is linked to B in D.

This shows (39.24), and hence that $M^* = L$. So the dual of a strict gammoid is a transversal matroid.

To see that each transversal matroid is the dual of a strict gammoid, we show that the construction described above can be reversed. Let L be the transversal matroid induced by the family $\mathcal{X} = (X_i \mid i = 1, \ldots, m)$ of sets. By (39.18) we can assume that $\mathcal{X}$ has a transversal. Hence we can assume that $i \in X_i$ for $i = 1, \ldots, m$ (by renaming). Let $V := X_1 \cup \cdots \cup X_m$ and let

(39.26) $A := \{(u, v) \mid v \in \{1, \ldots, m\}, u \in X_v\}$.

Let $D = (V, A)$ and define $U := V \setminus \{1, \ldots, m\}$. Since D, U and $\mathcal{X}$ are related as in (39.23), we again have (39.25). So L is equal to the dual of the strict gammoid induced by D, U. ∎

This theorem has a number of implications for the interrelations of the classes of transversal matroids and of gammoids. Consider the following class of matroids, introduced by Ingleton and Piff [1973]. Let $G = (V, E)$ be a bipartite graph, with colour classes U and W. Let $M = (V, \mathcal{I})$ be the transversal matroid induced by the family $(\{v\} \cup N(v) \mid v \in U)$ (where $N(v)$ is the set of neighbours of v). So $B \subseteq V$ is a base of M if and only if $(U \setminus B) \cup (W \cap B)$ is matchable in G (that is, it induces a subgraph of G having a perfect matching).

Any such matroid M is called a *deltoid* (*induced by* G, U, W). Then M^* is the deltoid induced by G, W, U. So

(39.27) the dual of a deltoid is a deltoid again.

Now

(39.28) transversal matroids are exactly those matroids that are the restriction of a deltoid.

Indeed, each deltoid is a transversal matroid, and hence the restriction of any deltoid is a transversal matroid (as the class of transversal matroids is closed under taking restrictions). Conversely, any transversal matroid, induced by (say) $X_1, \ldots, X_m$ is the restriction to W of the deltoid induced by the bipartite graph G with colour classes $U := \{1,\ldots,m\}$ and $W := X_1 \cup \cdots \cup X_m$, with $i \in U$ and $x \in W$ adjacent if and only if $x \in X_i$. (Assuming without loss of generality that $U \cap W = \emptyset$.) This shows (39.28).

Then (39.27) and (39.28) give with Theorem 39.5:

(39.29) the strict gammoids are exactly the contractions of the deltoids.

Indeed, the strict gammoids are the duals of transversal matroids, hence the duals of restrictions of deltoids, and therefore the contractions of (the duals of) deltoids.

This gives:

Corollary 39.5a. *The gammoids are exactly the contractions of the transversal matroids.*

Proof. Gammoids are the restrictions of strict gammoids, hence the restrictions of contractions of deltoids, hence the contractions of restrictions of deltoids, therefore the contractions of transversal matroids. ∎

Similarly:

(39.30) the gammoids are exactly the minors of deltoids,

which implies (with (39.27)) a result of Mason [1972]:

(39.31) the class of gammoids is closed under taking minors and duals.

Theorem 39.5 also implies, with (39.20), that gammoids are representable over all fields, except for a finite number of finite fields (Mason [1972]). In fact, Lindström [1973] showed that any gammoid $(S,\mathcal{I})$ is representable over each field with at least $2^{|S|}$ elements.

Edmonds and Fulkerson [1965] showed that one gets a transversal matroid as follows. Let $G = (V,E)$ be an undirected graph and let $S \subseteq V$. Let $\mathcal{I}$ be the collection of subsets of S which are covered by some matching in G. Then $M = (S,\mathcal{I})$ is a matroid (which is easy to show), called the *matching matroid* of G. In fact, any matching matroid is a transversal matroid. To prove this, we may assume $S = V$. Let $D(G)$, $A(G)$, $C(G)$ form the Edmonds-Gallai decomposition of G (Section 24.4b). Let $\mathcal{K}$ be the collection of components of $G[D(G)]$. Let $\mathcal{X}$ be the family of sets

$$
(39.32)\qquad \begin{array}{ll}
\{v\} & \text{for each } v \in A(G) \cup C(G), \\
N(v) \cap D(G) & \text{for each } v \in A(G), \\
K, \text{ repeated } |K|-1 \text{ times}, & \text{for each } K \in \mathcal{K}.
\end{array}
$$

Then M is equal to the transversal matroid induced by $\mathcal{X}$, as is easy to derive from the properties of the Edmonds-Gallai decomposition. A min-max relation for the rank function is given by Theorem 24.6.

It is straightforward to see that, conversely, each transversal matroid is a matching matroid, by taking G bipartite.

39.5. Characterizing matroids by bases

In Section 39.1, the notion of matroid is defined by 'axioms' in terms of the independent sets. There are several other axiom systems that characterize matroids. In this and the next sections we give a number of them.

Clearly, a matroid is determined by the collection of its bases, since a set is independent if and only if it is contained in a base. Conditions characterizing a collection of bases of a matroid are given in the following theorem (Whitney [1935]).

Theorem 39.6. *Let S be a set and let $\mathcal{B}$ be a nonempty collection of subsets of S. Then the following are equivalent:*

(39.33) (i) *$\mathcal{B}$ is the collection of bases of a matroid;*
(ii) *if $B, B' \in \mathcal{B}$ and $x \in B' \setminus B$, then $B' - x + y \in \mathcal{B}$ for some $y \in B \setminus B'$;*
(iii) *if $B, B' \in \mathcal{B}$ and $x \in B' \setminus B$, then $B - y + x \in \mathcal{B}$ for some $y \in B \setminus B'$.*

Proof. (i)⇒(ii): Let $\mathcal{B}$ be the collection of bases of a matroid $(S, \mathcal{I})$. Then all sets in $\mathcal{B}$ have the same size. Now let $B, B' \in \mathcal{B}$ and $x \in B' \setminus B$. Since $B' - x \in \mathcal{I}$, there exists a $y \in B \setminus B'$ with $B'' := B' - x + y \in \mathcal{I}$. Since $|B''| = |B'|$, we know $B'' \in \mathcal{B}$.

(iii)⇒(i): (iii) directly implies that no set in $\mathcal{B}$ is contained in another. Let $\mathcal{I}$ be the collection of sets I with $I \subseteq B$ for some $B \in \mathcal{B}$. We check (39.3). Let $I, J \in \mathcal{I}$ with $|I \setminus J| = 1$ and $|J \setminus I| = 2$. Let $I \setminus J = \{x\}$.

Consider sets $B, B' \in \mathcal{B}$ with $I \subseteq B$, $J \subseteq B'$. If $x \in B'$, we are done. So assume $x \notin B'$. Then by (iii), $B' - y + x \in \mathcal{B}$ for some $y \in B' \setminus B$. As $|J \setminus I| = 2$, there is a $z \in J \setminus I$ with $z \neq y$. Then $I + z \subseteq B' - y + x$, and so $I + z \in \mathcal{I}$.

(ii)⇒(iii): By the foregoing we know that (iii) implies (ii). Now axioms (ii) and (iii) interchange if we replace $\mathcal{B}$ by the collection of complements of sets in $\mathcal{B}$. Hence also the implication (ii)⇒(iii) holds. ∎

The equivalence of (ii) and (iii) also follows from the fact that the collection of complements of bases of a matroid is the collection of bases of the dual matroid. Conversely, Theorem 39.6 implies that the dual indeed is a matroid.

39.6. Characterizing matroids by circuits

A matroid is determined by the collection of its circuits, since a set is independent if and only if it contains no circuit. Conditions characterizing a collection of circuits of a matroid are given in the following theorem (Whitney

[1935] proved (i)$\Leftrightarrow$(iii), and Robertson and Weston [1958] (and also Lehman [1964] and Asche [1966]) proved (i)$\Leftrightarrow$(ii)).

Theorem 39.7. *Let S be a set and let $\mathcal{C}$ be a collection of nonempty subsets of S, such that no two sets in $\mathcal{C}$ are contained in each other. Then the following are equivalent:*

(39.34) (i) *$\mathcal{C}$ is the collection of circuits of a matroid;*
(ii) *if $C, C' \in \mathcal{C}$ with $C \neq C'$ and $x \in C \cap C'$, then $(C \cup C') \setminus \{x\}$ contains a set in $\mathcal{C}$;*
(iii) *if $C, C' \in \mathcal{C}$, $x \in C \cap C'$, and $y \in C \setminus C'$, then $(C \cup C') \setminus \{x\}$ contains a set in $\mathcal{C}$ containing y.*

Proof. (i)$\Rightarrow$(iii): Let $\mathcal{C}$ be the collection of circuits of a matroid $(S, \mathcal{I})$ and let $\mathcal{B}$ be its collection of bases. Let $C, C' \in \mathcal{C}$, $x \in C \cap C'$, and $y \in C \setminus C'$. We can assume that $S = C \cup C'$. Let $B, B' \in \mathcal{B}$ with $B \supseteq C - y$ and $B' \supseteq C' - x$. Then $y \notin B$ and $x \notin B'$ (since $C \not\subseteq B$ and $C' \not\subseteq B'$).

We can assume that $y \notin B'$. Otherwise, $y \in B' \setminus B$, and hence by (ii) of Theorem 39.6, there exists a $z \in B \setminus B'$ with $B'' := B' - y + z \in \mathcal{B}$. Then $z \neq x$, since otherwise $C' \subseteq B''$. Hence, replacing B' by B'' gives $y \notin B'$.

As $y \notin B'$, we know $B' \cup \{y\} \notin \mathcal{I}$, and hence there exists a $C'' \in \mathcal{C}$ contained in $B' \cup \{y\}$. As $C'' \not\subseteq B'$, we know $y \in C''$. Moreover, as $x \notin B'$ we know $x \notin C''$.

(iii)$\Rightarrow$(ii): is trivial.

(ii)$\Rightarrow$(i): Let $\mathcal{I}$ be the collection of sets containing no set in $\mathcal{C}$ as a subset. We check (39.3). Let $I, J \in \mathcal{I}$ with $|I \setminus J| = 1$ and $|J \setminus I| = 2$. Assume that $I + z \notin \mathcal{I}$ for each $z \in J \setminus I$. Let y be the element of $I \setminus J$. If $J + y \in \mathcal{I}$, then $I \cup J \in \mathcal{I}$, contradicting our assumption. So $J + y$ contains a set $C \in \mathcal{C}$. Then C is the unique set in $\mathcal{C}$ contained in $J + y$. For suppose that there is another, C' say. Again, $y \in C'$, and hence by (39.34)(ii) there exists a $C'' \in \mathcal{C}$ contained in $(C \cup C') \setminus \{y\}$. But then $C'' \subseteq J$, a contradiction.

As $C \not\subseteq I$, C intersects $J \setminus I$. Choose $x \in C \cap (J \setminus I)$. Then $X := J + y - x$ contains no set in $\mathcal{C}$ (as C is the only set in $\mathcal{C}$ contained in $J + y$). So $X \in \mathcal{I}$, implying that $I + z \in \mathcal{I}$ for the $z \in J \setminus I$ with $z \neq x$. ∎

This theorem implies the following important property for a matroid $M = (S, \mathcal{I})$:

(39.35) for any independent set I and any $s \in S \setminus I$ there is at most one circuit contained in $I \cup \{s\}$.

39.6a. A characterization of Lehman

Lehman [1964] showed that the cocircuits of a matroid M are exactly the inclusionwise minimal nonempty subsets D of S with $|D \cap C| \neq 1$ for each circuit C of M.

To show this, it suffices to show that

(39.36) (i) $|D \cap C| \neq 1$ for each cocircuit D and circuit C,
(ii) for each nonempty $D \subseteq S$, if $|D \cap C| \neq 1$ for each circuit C, then D contains a cocircuit; that is, then D is dependent in M^*.

To see (i), suppose that $D \cap C = \{s\}$ for some circuit C and cocircuit D. As $D - s$ is independent in M^*, M has a base B disjoint from $D - s$. Since $C - s$ is disjoint from $D - s$ and since $C - s \in \mathcal{I}$, we can assume that $C - s \subseteq B$. Then $s \notin B$, and so B is disjoint from D. This implies that D is independent in M^*, contradicting the fact that D is a circuit in M^*. This shows (i).

To see (ii), let $\emptyset \neq D \subseteq S$ with $|D \cap C| \neq 1$ for each circuit C. We show that D is dependent in M^*. Suppose not. Then M has a base B disjoint from D. Choose $s \in D$. Then $B + s$ contains a circuit C with $s \in C$. Hence $D \cap C = \{s\}$, contradicting our assumption, thus showing (ii).

39.7. Characterizing matroids by rank functions

The *rank function* of a matroid $M = (S, \mathcal{I})$ is the function $r_M : \mathcal{P}(S) \to \mathbb{Z}_+$ given by:

(39.37) $$r_M(U) := \max\{|Z| \mid Z \in \mathcal{I}, Z \subseteq U\}$$

for $U \subseteq S$. Again, a matroid is determined by its rank function, as a set U is independent if and only if $r(U) = |U|$. Conditions characterizing a rank function are given by the following theorem (Whitney [1935]; necessity was also shown (in a different terminology) by Bergmann [1929] and Nakasawa [1935]):

Theorem 39.8. *Let S be a set and let $r : \mathcal{P}(S) \to \mathbb{Z}_+$. Then r is the rank function of a matroid if and only if for all $T, U \subseteq S$:*

(39.38) (i) $r(T) \leq r(U) \leq |U|$ *if* $T \subseteq U$,
(ii) $r(T \cap U) + r(T \cup U) \leq r(T) + r(U)$.

Proof. *Necessity.* Let r be the rank function of a matroid $(S, \mathcal{I})$. Choose $T, U \subseteq S$. Clearly (39.38)(i) holds. To see (ii), let I be an inclusionwise maximal set in $\mathcal{I}$ with $I \subseteq T \cap U$ and let J be an inclusionwise maximal set in $\mathcal{I}$ with $I \subseteq J \subseteq T \cup U$. Since $(S, \mathcal{I})$ is a matroid, we know that $r(T \cap U) = |I|$ and $r(T \cup U) = |J|$. Then

(39.39) $$r(T) + r(U) \geq |J \cap T| + |J \cap U| = |J \cap (T \cap U)| + |J \cap (T \cup U)| \geq |I| + |J| = r(T \cap U) + r(T \cup U);$$

that is, we have (39.38)(ii).

Sufficiency. Let $\mathcal{I}$ be the collection of subsets I of S with $r(I) = |I|$. We show that $(S, \mathcal{I})$ is a matroid, with rank function r.

Trivially, $\emptyset \in \mathcal{I}$. Moreover, if $I \in \mathcal{I}$ and $J \subseteq I$, then

(39.40) $\qquad r(J) \geq r(I) - r(I \setminus J) \geq |I| - |I \setminus J| = |J|.$

So $J \in \mathcal{I}$.

In order to check (39.3), let $I, J \in \mathcal{I}$ with $|I \setminus J| = 1$ and $|J \setminus I| = 2$. Let $J \setminus I = \{z_1, z_2\}$. If $I + z_1, I + z_2 \notin \mathcal{I}$, we have $r(I + z_1) = r(I + z_2) = |I|$. Then by (39.38)(ii),

(39.41) $\qquad r(J) \leq r(I + z_1 + z_2) \leq r(I + z_1) + r(I + z_2) - r(I) = |I| < |J|,$

contradicting the fact that $J \in \mathcal{I}$.

So $(S, \mathcal{I})$ is a matroid. Its rank function is r, since $r(U) = \max\{|I| \mid I \subseteq U, I \in \mathcal{I}\}$ for each $U \subseteq S$. Here $\geq$ follows from (39.38)(i), since if $I \subseteq U$ and $I \in \mathcal{I}$, then $r(U) \geq r(I) = |I|$. Equality can be shown by induction on $|U|$, the case $U = \emptyset$ being trivial. If $U \neq \emptyset$, choose $y \in U$. By induction, there is an $I \subseteq U - y$ with $I \in \mathcal{I}$ and $|I| = r(U - y)$. If $r(U) = r(U - y)$ we are done, so assume $r(U) > r(U - y)$. Then $I + y \in \mathcal{I}$, since $r(I + y) \geq r(I) + r(U) - r(U - y) \geq |I| + 1$. Moreover, $r(U) \leq r(U - y) + r(\{y\}) \leq |I| + 1$. This proves equality for U. ∎

Set functions satisfying condition (39.38)(ii) are called *submodular*, and will be studied in Chapter 44.

Whitney [1935] also showed that (39.38) is equivalent to:

(39.42)
(i) $r(\emptyset) = 0$,
(ii) $r(U) \leq r(U + s) \leq r(U) + 1$ for $U \subseteq S$, $s \in S \setminus U$,
(iii) for all $U \subseteq S$, $s, t \in S \setminus U$, if $r(U + s) = r(U + t) = r(U)$, then $r(U + s + t) = r(U)$.

The proof above in fact uses only these properties of r.

The following equivalent form of Theorem 39.8 will be useful.

Corollary 39.8a. *Let S be a finite set and let $\mathcal{I}$ be a nonempty collection of subsets of S, closed under taking subsets. For $U \subseteq S$, let $r(U)$ be the maximum size of a subset of U that belongs to $\mathcal{I}$. Then $(S, \mathcal{I})$ is a matroid if and only if r satisfies* (39.38)(ii) *for all $T, U \subseteq S$.*

Proof. Necessity follows directly from Theorem 39.8. To see sufficiency, it is easy to see that r satisfies (39.38)(i). So by Theorem 39.8, r is the rank function of some matroid $M = (S, \mathcal{J})$. Now: $I \in \mathcal{J} \iff r(I) = |I| \iff I \in \mathcal{I}$. Hence $\mathcal{I} = \mathcal{J}$, and so $(S, \mathcal{I})$ is a matroid. ∎

Note that if we can test in polynomial time if a given set is independent, we can also test in polynomial time if a given set is a base, or a circuit, and we can determine the rank of a given set in polynomial time.

39.8. The span function and flats

With any matroid $M = (S, \mathcal{I})$ we can define the *span function* $\text{span}_M : \mathcal{P}(S) \to \mathcal{P}(S)$ as follows:

(39.43) $$\text{span}_M(T) := \{s \in S \mid r_M(T \cup \{s\}) = r_M(T)\}$$

for $T \subseteq S$. If the matroid M is clear from the context, we write $\text{span}(T)$ for $\text{span}_M(T)$. Note that $T \subseteq \text{span}_M(T)$ and that

(39.44) $$r_M(\text{span}_M(T)) = r_M(T).$$

This follows directly from the fact that if $r_M(Y) > r_M(T)$, then $r_M(T \cup \{s\}) > r_M(T)$ for some $s \in Y$.

Note also that

(39.45) $$T \text{ is spanning} \iff \text{span}_M(T) = S$$

for any $T \subseteq S$. To see $\Longrightarrow$, let T be spanning. Then for each $s \in T$: $r_M(T + s) \leq r_M(S) = r_M(T)$. To see $\Longleftarrow$, suppose $\text{span}_M(T) = S$. Then $r_M(T) = r_M(\text{span}_M(T)) = r_M(S)$.

A *flat* in a matroid $M = (S, \mathcal{I})$ is a subset F of S with $\text{span}_M(F) = F$. A matroid is determined by its collection of flats, as is shown by:

(39.46) a subset I of S is independent if and only if for each $y \in I$ there is a flat F with $I - y \subseteq F$ and $y \notin F$.

Indeed, if I is independent and $y \in I$, let $F := \text{span}_M(I - y)$. Then F is a flat containing $I - y$, but not y, since $r_M(F + y) \geq r_M(I) > r_M(I - y) = r_M(F)$. Conversely, if I is not independent, then $y \in \text{span}_M(I - y)$ for some $y \in I$, and hence each flat containing $I - y$ also contains y.

39.8a. Characterizing matroids by span functions

It was observed by Mac Lane [1938] that the following characterizes span functions of matroids (sufficiency was shown by van der Waerden [1937]).

Theorem 39.9. *Let S be a finite set. A function $\text{span} : \mathcal{P}(S) \to \mathcal{P}(S)$ is the span function of a matroid if and only if:*

(39.47)
(i) *if $T \subseteq S$, then $T \subseteq \text{span}(T)$;*
(ii) *if $T, U \subseteq S$ and $U \subseteq \text{span}(T)$, then $\text{span}(U) \subseteq \text{span}(T)$;*
(iii) *if $T \subseteq S$, $t \in S \setminus T$, and $s \in \text{span}(T + t) \setminus \text{span}(T)$, then $t \in \text{span}(T + s)$.*

Proof. *Necessity.* Let span be the span function of a matroid $M = (S, \mathcal{I})$ with rank function r. Clearly, (39.47)(i) is satisfied. To see (39.47)(ii), let $U \subseteq \text{span}(T)$ and $s \in \text{span}(U)$. We show $s \in \text{span}(T)$. We can assume $s \notin T$. Then, by the submodularity of r,

(39.48) $$\begin{aligned} r(T \cup \{s\}) &\leq r(T \cup U \cup \{s\}) \leq r(T \cup U) + r(U \cup \{s\}) - r(U) \\ &= r(T \cup U) = r(T). \end{aligned}$$

(The last equality follows from (39.44).) This shows that $s \in \text{span}(T)$.

To see (39.47)(iii), note that $s \in \text{span}(T+t) \setminus \text{span}(T)$ is equivalent to: $r(T+t+s) = r(T+t)$ and $r(T+s) > r(T)$. Hence

(39.49) $$r(T+t+s) = r(T+t) \leq r(T) + 1 \leq r(T+s),$$

that is, $t \in \text{span}(T+s)$. This shows necessity of the conditions (39.47).

Sufficiency. Let a function span satisfy (39.47), and define

(39.50) $$\mathcal{I} := \{I \subseteq S \mid s \notin \text{span}(I-s) \text{ for each } s \in I\}.$$

We first show the following:

(39.51) $$\text{if } I \in \mathcal{I}, \text{ then } \text{span}(I) = I \cup \{t \mid I+t \notin \mathcal{I}\}.$$

Indeed, if $t \in \text{span}(I) \setminus I$, then $I+t \notin \mathcal{I}$, by definition of $\mathcal{I}$. Conversely, $I \subseteq \text{span}(I)$ by (39.47)(i). Moreover, if $I+t \notin \mathcal{I}$, then by definition of $\mathcal{I}$, $s \in \text{span}(I+t-s)$ for some $s \in I+t$. If $s=t$, then $t \in \text{span}(I)$ and we are done. So assume $s \neq t$; that is, $s \in I$. As $I \in \mathcal{I}$, we know that $s \notin \text{span}(I-s)$. So by (39.47)(iii) (for $T := I-s$), $t \in \text{span}(I)$, proving (39.51).

We now show that $M = (S,\mathcal{I})$ is a matroid. Trivially, $\emptyset \in \mathcal{I}$. To see that $\mathcal{I}$ is closed under taking subsets, let $I \in \mathcal{I}$ and $J \subseteq I$. We show that $J \in \mathcal{I}$. Suppose to the contrary that $s \in \text{span}(J-s)$ for some $s \in J$. By (39.47)(ii), $\text{span}(J-s) \subseteq \text{span}(I-s)$. Hence $s \in \text{span}(I-s)$, contradicting the fact that $I \in \mathcal{I}$.

In order to check (39.3), let $I, J \in \mathcal{I}$ with $|I \setminus J| = 1$ and $|J \setminus I| = 2$. Let $I \setminus J = \{i\}$ and $J \setminus I = \{j_1, j_2\}$. Assume that $I + j_1 \notin \mathcal{I}$. That is, $J + i - j_2 \notin \mathcal{I}$, and so, by (39.51) applied to $J - j_2$, $i \in \text{span}(J - j_2)$. Therefore, $I \subseteq \text{span}(J - j_2)$, and so $\text{span}(I) \subseteq \text{span}(J - j_2)$. So $j_2 \notin \text{span}(I)$ (as $J \in \mathcal{I}$), and therefore, by (39.51) applied to I, $I + j_2 \in \mathcal{I}$.

So M is a matroid. We finally show that span $= \text{span}_M$. Choose $T \subseteq S$. To see that $\text{span}(T) = \text{span}_M(T)$, let I be a base of T (in M). Then (using (39.51)),

(39.52) $$\text{span}_M(T) = I \cup \{x \mid I + x \notin \mathcal{I}\} = \text{span}(I) \subseteq \text{span}(T).$$

So we are done by showing $\text{span}(T) \subseteq \text{span}(I)$; that is, by (39.47)(ii), $T \subseteq \text{span}(I)$. Choose $t \in T \setminus I$. By the maximality of I, we know $I + t \notin \mathcal{I}$, and hence, by (39.51), $t \in \text{span}(I)$. ∎

39.8b. Characterizing matroids by flats

Conditions characterizing collections of flats of a matroid are given in the following theorem (Bergmann [1929]):

Theorem 39.10. *Let S be a set and let $\mathcal{F}$ be a collection of subsets of S. Then $\mathcal{F}$ is the collection of flats of a matroid if and only if:*

(39.53)
- (i) *$S \in \mathcal{F}$;*
- (ii) *if $F_1, F_2 \in \mathcal{F}$, then $F_1 \cap F_2 \in \mathcal{F}$;*
- (iii) *if $F \in \mathcal{F}$ and $t \in S \setminus F$, and F' is the smallest flat containing $F+t$, then there is no flat F'' with $F \subset F'' \subset F'$.*

Proof. *Necessity.* Let $\mathcal{F}$ be the collection of flats of a matroid $M = (S, \mathcal{I})$. Condition (39.53)(i) is trivial, and condition (39.53)(ii) follows from $\mathrm{span}_M(F_1 \cap F_2) \subseteq \mathrm{span}_M(F_1) \cap \mathrm{span}_M(F_2) = F_1 \cap F_2$. To see (39.53)(iii), suppose that such an F'' exists. Choose $s \in F'' \setminus F$. So $s \notin \mathrm{span}_M(F)$. As $F' \not\subseteq F''$, we have $t \notin \mathrm{span}_M(F+s)$. Therefore, by (39.47)(iii) for $T := F$, $s \notin \mathrm{span}_M(F) = F'$, a contradiction.

Sufficiency. Let $\mathcal{F}$ satisfy (39.53). For $Y \subseteq S$, let $\mathrm{span}(Y)$ be the smallest set in $\mathcal{F}$ containing Y. Since $F \in \mathcal{F} \iff \mathrm{span}(F) = F$, it suffices to show that span satisfies the conditions (39.47). Here (39.47)(i) and (ii) are trivial. To see (39.47)(iii), let $T \subseteq S$, $t \in S \setminus T$, and $s \in \mathrm{span}(T+t) \setminus \mathrm{span}(T)$. Then $\mathrm{span}(T) \subset \mathrm{span}(T+s) \subseteq \mathrm{span}(T+t)$. Hence, by (39.53)(iii), $\mathrm{span}(T+s) = \mathrm{span}(T+t)$, and hence $t \in \mathrm{span}(T+s)$. ∎

39.8c. Characterizing matroids in terms of lattices

Bergmann [1929] and Birkhoff [1935a] characterized matroids in terms of lattices. A partially ordered set $(L, \leq)$ is called a *lattice* if

(39.54)
(i) for all $A, B \in L$ there is a unique element, called $A \wedge B$, satisfying $A \wedge B \leq A, B$ and $C \leq A \wedge B$ for all $C \leq A, B$;
(ii) for all $A, B \in L$ there is a unique element, called $A \vee B$, satisfying $A \vee B \geq A, B$ and $C \geq A \vee B$ for all $C \geq A, B$.

$A \wedge B$ and $A \vee B$ are called the *meet* and *join* respectively of A and B. Here we assume lattices to be finite. Then a lattice has a unique minimal element, denoted by 0. The *rank* of an element A is the maximum number n of elements $x_1, \ldots, x_n$ with $0 < x_1 < \cdots < x_n = A$. An element of rank 1 is called a *point* or *atom*.

Call a lattice a *point lattice* if each element is a join of points, and a *matroid lattice* (or a *geometric lattice*) if it is isomorphic to the lattice of flats of a matroid. Trivially, each matroid lattice is a point lattice. Moreover, a matroid without loops and parallel elements is completely determined by the lattice of flats.

In the following theorem, the equivalence of (i) and (ii), and the implication (ii)⇒(iv) are due (in a different terminology) to Bergmann [1929]; the equivalence of (iii) and (iv) was shown by Birkhoff [1933], and the implication (iii)⇒(i) was shown by Birkhoff [1935a].

In a partially ordered set $(L, \leq)$ an element y is said to *cover* an element x if $x < y$ and there is no z with $x < z < y$.

Theorem 39.11. *For any finite point lattice $(L, \leq)$, with rank function r, the following are equivalent:*

(39.55)
(i) *L is a matroid lattice;*
(ii) *for each $a \in L$ and each point p, if $p \not\leq a$, then $a \vee p$ covers a;*
(iii) *for each $a, b \in L$, if a and b cover $a \wedge b$, then $a \vee b$ covers a and b;*
(iv) *$r(a) + r(b) \geq r(a \vee b) + r(a \wedge b)$ for all $a, b \in L$.*

Proof. (i)⇒(iv): Let L be the lattice of flats of a matroid $M = (S, \mathcal{I})$, with rank function r_M. We can assume that M has no loops and no parallel elements. Then for any flat F we have $r(F) = r_M(F)$, since $r_M(F)$ is equal to the maximum number k of nonempty flats $F_1 \subset \cdots \subset F_k$ with $F_k = F$. So (iv) follows from Theorem 39.8.

(iv)⇒(iii): We first show that (iv) implies that if b covers a, then $r(b) = r(a)+1$. As b is a join of points, and as b covers a, we know that $b = a \vee p$ for some point p with $p \not\leq a$. Hence $r(b) = r(a \vee p) \leq r(a) + r(p) - r(a \wedge p) = r(a) + r(p) - r(0) = r(a) + 1$. As $r(b) > r(a)$, we have $r(b) = r(a) + 1$.

To derive (iii) from (iv), let a and b cover $a \wedge b$. Then $r(a) = r(b) = r(a \wedge b) + 1$. Hence $r(a \vee b) \leq r(a) + r(b) - r(a \wedge b) = r(a) + 1$. Hence $a \vee b$ covers a. Similarly, $a \vee b$ covers b.

(iii)⇒(ii): We derive (ii) from (iii) by induction on $r(a)$. If $a = 0$, the statement is trivial. If $a > 0$, let a' be an element covered by a. Then, by induction, $a' \vee p$ covers a'. So $a' = a \wedge (a' \vee p)$. Hence by (iii), $a \vee (a' \vee p) = a \vee p$ covers a.

(ii)⇒(i): Let S be the set of points of L, and for $f \in L$ define $F_f := \{s \in S \mid s \leq f\}$. Let $\mathcal{F} := \{F_f \mid f \in L\}$. Then for all $f_1, f_2 \in L$ we have:

(39.56) $$f_1 \leq f_2 \iff F_{f_1} \subseteq F_{f_2}.$$

Here $\Longrightarrow$ is trivial, while $\Longleftarrow$ follows from the fact that for each $f \in L$ we have $f = \bigvee F_f$, as L is a point lattice.

By (39.56), $(L, \leq)$ is isomorphic to $(\mathcal{F}, \subseteq)$. Moreover, by (39.54)(i), $F_{f_1 \wedge f_2} = F_{f_1} \cap F_{f_2}$. So $\mathcal{F}$ is closed under intersections, implying (39.53)(ii), while (39.53)(i) is trivial. Finally, (39.53)(iii) follows from (39.55)(ii). ∎

Lattices satisfying (39.55)(iii) are called *upper semimodular*.

39.9. Further exchange properties

In this section we prove a number of exchange properties of bases, as a preparation to the forthcoming sections on matroid intersection algorithms.

An exchange property of bases, stronger than given in Theorem 39.6, is (Brualdi [1969c]):

Theorem 39.12. *Let $M = (S, \mathcal{I})$ be a matroid. Let B_1 and B_2 be bases and let $x \in B_1 \setminus B_2$. Then there exists a $y \in B_2 \setminus B_1$ such that both $B_1 - x + y$ and $B_2 - y + x$ are bases.*

Proof. Let C be the unique circuit in $B_2 + x$ (cf. (39.35)). Then $(B_1 \cup C) - x$ is spanning, since $x \in \text{span}_M(C - x) \subseteq \text{span}_M((B_1 \cup C) - x)$, implying $\text{span}((B_1 \cup C) - x) = \text{span}(B_1 \cup C) = S$.

Hence there is a base B_3 with $B_1 - x \subseteq B_3 \subseteq (B_1 \cup C) - x$. So $B_3 = B_1 - x + y$ for some y in $C - x$. Therefore, $B_2 - y + x$ is a base, as it contains no circuit (since C is the only circuit in $B_2 + x$). ∎

Let $M = (S, \mathcal{I})$ be a matroid. For any $I \in \mathcal{I}$ define the (bipartite) directed graph $D_M(I) = (S, A_M(I))$, or briefly $(S, A(I))$, by:

(39.57) $$A(I) := \{(y, z) \mid y \in I, z \in S \setminus I, I - y + z \in \mathcal{I}\}.$$

Repeated application of the exchange property described in Theorem 39.12 gives (Brualdi [1969c]):

Corollary 39.12a. *Let $M = (S,\mathcal{I})$ be a matroid and let $I, J \in \mathcal{I}$ with $|I| = |J|$. Then $A(I)$ contains a perfect matching on $I \triangle J$.*[1]

Proof. By truncating M, we can assume that I and J are bases of M. We prove the lemma by induction on $|I \setminus J|$. We can assume $|I \setminus J| \geq 1$. Choose $y \in I \setminus J$. By Theorem 39.12, $I - y + z \in \mathcal{I}$ and $J - z + y \in \mathcal{I}$ for some $z \in J \setminus I$. By induction, applied to I and $J' := J - z + y$, $A(I)$ has a perfect matching N on $I \triangle J'$. Then $N \cup \{(y, z)\}$ is a perfect matching on $I \triangle J$. ∎

Corollary 39.12a implies the following characterization of maximum-weight bases:

Corollary 39.12b. *Let $M = (S,\mathcal{I})$ be a matroid, let B be a base of M, and let $w : S \to \mathbb{R}$ be a weight function. Then B is a base of maximum weight $\iff$ $w(B') \leq w(B)$ for every base B' with $|B' \setminus B| = 1$.*

Proof. Necessity being trivial, we show sufficiency. Suppose to the contrary that there is a base B' with $w(B') > w(B)$. Let N be a perfect matching in $A(B)$ covering $B \triangle B'$. As $w(B') > w(B)$, there is an edge (y, z) in N with $w(z) > w(y)$, where $y \in B \setminus B'$ and $z \in B' \setminus B$. Hence $w(B - y + z) > w(B)$, contradicting the condition. ∎

The following forms a counterpart to Corollary 39.12a (Krogdahl [1974, 1976,1977]):

Theorem 39.13. *Let $M = (S,\mathcal{I})$ be a matroid and let $I \in \mathcal{I}$. Let $J \subseteq S$ be such that $|I| = |J|$ and such that $A(I)$ contains a unique perfect matching N on $I \triangle J$. Then J belongs to $\mathcal{I}$.*

Proof. Since N is unique, we can order N as $(y_1, z_1), \ldots, (y_t, z_t)$ such that $(y_i, z_j) \notin A(I)$ if $1 \leq i < j \leq t$. Suppose that $J \notin \mathcal{I}$, and let C be a circuit contained in J. Choose the smallest i with $z_i \in C$. Then $(y_i, z) \notin A(I)$ for all $z \in C - z_i$ (since $z = z_j$ for some $j > i$). Therefore, $z \in \mathrm{span}(I - y_i)$ for all $z \in C - z_i$. So $C - z_i \subseteq \mathrm{span}(I - y_i)$, and therefore $z_i \in C \subseteq \mathrm{span}(C - z_i) \subseteq \mathrm{span}(I - y_i)$, contradicting the fact that $I - y_i + z_i$ is independent. ∎

This implies:

Corollary 39.13a. *Let $M = (S,\mathcal{I})$ be a matroid and let $I \in \mathcal{I}$. Let $J \subseteq S$ be such that $|I| = |J|$ and $r_M(I \cup J) = |I|$, and such that $A(I)$ contains a unique perfect matching N on $I \triangle J$. Let $s \notin I \cup J$ with $I + s \in \mathcal{I}$. Then $J + s \in \mathcal{I}$.*

Proof. Let t be a new element and let $M' = (S \cup \{t\}, \mathcal{I}')$ be the matroid with $F \in \mathcal{I}'$ if and only if $F \setminus \{t\} \in \mathcal{I}$. Then $N' := N \cup \{(t, s)\}$ forms a

[1] A *perfect matching on* a vertex set U in a digraph is a set of vertex-disjoint arcs such that U is the set of tails and heads of these arcs.

unique perfect matching on $(I \triangle J) \cup \{s,t\}$ in $D_{M'}(I \cup \{t\})$ (since there is no arc from t to $J \setminus I$, as $I + j \notin \mathcal{I}$ for all $j \in J \setminus I$, since $r_M(I \cup J) = |I|$). So by Theorem 39.13, $J \cup \{s\}$ is independent in M', and hence in M. ∎

39.9a. Further properties of bases

Bases satisfy the following exchange property, stronger than that described in Theorem 39.12 (conjectured by G.-C. Rota, and proved by Brylawski [1973], Greene [1973], Woodall [1974a]):

(39.58) if B_1 and B_2 are bases and B_1 is partitioned into X_1 and Y_1, then B_2 can be partitioned into X_2 and Y_2 such that $X_1 \cup Y_2$ and $Y_1 \cup X_2$ are bases.

This will be proved in Section 42.1a (using the matroid union theorem).

Other exchange properties of bases were given by Greene [1974a] and Kung [1978a]. Decomposing exchanges was studied by Gabow [1976b].

In Schrijver [1979c] it was shown that the exchange property described in Corollary 16.8b for bipartite graphs and, more generally, in Theorem 9.12 for directed graphs, in fact characterizes systems that correspond to matroids.

To this end, let U and W be disjoint sets and let Λ be a collection of pairs (X,Y) with $X \subseteq U$ and $Y \subseteq W$. Call (U,W,Λ) a *bimatroid* (or *linking system*) if:

(39.59)
(i) $(\emptyset,\emptyset) \in \Lambda$;
(ii) if $(X,Y) \in \Lambda$ and $x \in X$, then $(X-x, Y-y) \in \Lambda$ for some $y \in Y$;
(iii) if $(X,Y) \in \Lambda$ and $y \in Y$, then $(X-x, Y-y) \in \Lambda$ for some $x \in X$;
(iv) if $(X_1,Y_1),(X_2,Y_2) \in \Lambda$, then there is an $(X,Y) \in \Lambda$ with $X_1 \subseteq X \subseteq X_1 \cup X_2$ and $Y_2 \subseteq Y \subseteq Y_1 \cup Y_2$.

Note that (ii) and (iii) imply that $|X| = |Y|$ for each $(X,Y) \in \Lambda$.

To describe the relation with matroids, define:

(39.60) $\mathcal{B} := \{(U \setminus X) \cup Y \mid (X,Y) \in \Lambda\}$.

So $\mathcal{B}$ determines Λ. Then (Schrijver [1979c]):

(39.61) (U,W,Λ) is a bimatroid if and only if $\mathcal{B}$ is the collection of bases of a matroid on $U \cup W$, with $U \in \mathcal{B}$.

So bimatroids are in one-to-one correspondence with pairs (M,B) of a matroid M and a base B of M, and the conditions (39.59) yield a characterization of matroids. An equivalent axiom system characterizing matroids was given by Kung [1978b].

(Bapat [1994] gave an extension of Kőnig's matching theorem to bimatroids.)

39.10. Further results and notes

39.10a. Further notes

Dilworth [1944] showed that if $r : \mathcal{P}(S) \to \mathbb{Z}$ satisfies (39.38) and $r(U) \geq 0$ if $U \neq \emptyset$, then

(39.62) $\mathcal{I} := \{I \subseteq S \mid \forall \text{ nonempty } U \subseteq I : |U| \leq r(U)\}$

is the collection of independent sets of a matroid M. Its rank function satisfies:

(39.63) $r_M(U) = \min(r(U_1) + \cdots + r(U_t))$,

where the minimum ranges over partitions of U into nonempty subsets $U_1, \ldots, U_t$ ($t \geq 0$). If $G = (V, E)$ is a graph, and we define $r(F) := \left|\bigcup F\right| - 1$ for $F \subseteq E$, we obtain the cycle matroid of G (this also was shown by Dilworth [1944]).[2]

Conforti and Laurent [1988] showed the following sharpening of Corollary 39.8a. Let $\mathcal{C}$ be a collection of subsets of a set S and let $f : \mathcal{C} \to \mathbb{Z}_+$. Let $\mathcal{I}$ be the collection of subsets T of S with $|T \cap U| \leq f(U)$ for each $U \in \mathcal{C}$. For $T \subseteq S$, let $r(T)$ be the maximum size of a subset of T that belongs to $\mathcal{I}$. Then $(S, \mathcal{I})$ is a matroid if and only if r satisfies the submodular inequality (39.38)(ii) for all $Y, Z \in \mathcal{C}$ with $Y \cap Z \neq \emptyset$. In fact, in the right-hand side of this inequality, r may be replaced by f.

Jensen and Korte [1982] showed that there is no polynomial-time algorithm to find the minimum size of a circuit of a matroid, if the matroid is given by an oracle for testing independence. For binary matroids (represented by binary vectors), the problem of finding a minimum-size circuit was shown by Vardy [1997] to be NP-complete (solving a problem of Berlekamp, McEliece, and van Tilborg [1978], who showed the NP-completeness of finding the minimum size of a circuit containing a given element of the matroid, and of finding a circuit of given size). If we know that a matroid is binary, a vector representation can be derived by a polynomially bounded number of calls from an independence testing oracle.

For further studies of the complexity of matroid properties, see Hausmann and Korte [1978], Robinson and Welsh [1980], and Jensen and Korte [1982].

Extensions of matroid theory to infinite structures were considered by Rado [1949a], Bleicher and Preston [1961], Johnson [1961], and Dlab [1962,1965].

Standard references on matroid theory are Welsh [1976] and Oxley [1992]. The book by Truemper [1992] focuses on decomposition of matroids. Earlier texts were given by Tutte [1965a,1971]. Elementary introductions to matroids were given by Wilson [1972b,1973], and a survey with applications to electrical networks and statics by Recski [1989]. Bixby [1982], Faigle [1987], Lee and Ryan [1992], and Bixby and Cunningham [1995] survey matroid optimization and algorithms. White [1986, 1987,1992] offers a collection of surveys on matroids, and Kung [1986] is a source book on matroids. Stern [1999] focuses on semimodular lattices. Books discussing matroid optimization include Lawler [1976b], Papadimitriou and Steiglitz [1982], Gondran and Minoux [1984], Nemhauser and Wolsey [1988], Parker and Rardin [1988], Cook, Cunningham, Pulleyblank, and Schrijver [1998], and Korte and Vygen [2000].

39.10b. Historical notes on matroids

The idea of a matroid, that is, of abstract dependence, seems to have been developed historically along a number of independent lines during the period 1900-1935. Independently, different axiom systems were given, each of which is equivalent to

[2] $\bigcup F$ denotes the union of the edges (as sets) in F.

that of a matroid. It indicates the naturalness of the concept. Only at the end of the 1930s a synthesis of the different streams was obtained.

There is a line, starting with the *Dualgruppen* (dual groups = lattices) of Dedekind [1897,1900], introduced in order to study modules (= additive subgroups) of numbers. They give rise to lattices satisfying what Dedekind called the *Modulgesetz* (module law). Later, independently, Birkhoff [1933] studied such lattices, calling them initially B-lattices, and later (after he had learned about Dedekind's earlier work), *modular lattices*. Both Dedekind and Birkhoff considered, in their studies of modular lattices, an auxiliary property that characterizes so-called *semimodular lattices*. If the lattice is a point lattice (that is, each element of the lattice is a join of atoms (points)), then such semimodular lattices are exactly the lattices of flats of a matroid. This connection was pointed out by Birkhoff [1935a] directly after Whitney's introduction of matroids.

A second line concerns exchange properties of bases. It starts with the new edition of the *Ausdehnungslehre* of Grassmann [1862], where he showed that each linearly independent set can be extended to a bases, using elements from a given base. Next Steinitz [1910], in his fundamental paper *Algebraische Theorie der Körper* (Algebraic Theory of Fields), showed that algebraic dependence has a number of basic properties, which makes it into a matroid (like the equicardinality of bases), and he derived some other properties from these basic properties (thus deriving essentially properties of matroids). In a subsequent paper, Steinitz [1913] gave, as an auxiliary result, the property that is now called *Steinitz' exchange property* for linearly independent sets of vectors. Steinitz did not mention the similarities to his earlier results on algebraic dependence. These similarities were observed by Haupt [1929a] and van der Waerden [1930] in their books on 'modern' algebra. They formulated properties shared by linear and algebraic dependence that are equivalent to matroids. In the second edition of his book, van der Waerden [1937] condensed these properties to three properties, and gave a unified treatment of linear and algebraic dependence. Mac Lane [1938] observed the relation of this work to the work on lattices and matroids.

A third line pursued the axiomatization of geometry, which clearly can be rooted back to as early as Euclid. At the beginning of the 20th century this was considered by, among others, Hilbert and Veblen. Bergmann [1929] aimed at giving a lattice-theoretical basis for affine geometry, and from lattice-theoretical conditions equivalent to matroids (cf. Theorem 39.11 above) he derived a number of properties, like the equicardinality of bases and the submodularity of the rank function. In their book *Grundlagen der Mathematik I* (Foundations of Mathematics I), Hilbert and Bernays [1934] gave axioms for the collinearity of triples of points, amounting to the fact that any two distinct points belong to exactly one line. A direct extension of these axioms to general dimensions gives the axioms described by Nakasawa [1935], that are again equivalent to the matroid axioms. He introduced the concept of a $\mathcal{B}_1$-space, equivalent to a matroid. In fact, the only reference in Nakasawa [1935] is to the book *Grundlagen der Elementargeometrie* (Foundations of Elementary Geometry) of Thomsen [1933], in which a different axiom system, the *Zyklenkalkül* (cycle calculus), was given (not equivalent to matroids). Nakasawa only gave subsets of linear spaces as an example. In a sequel to his paper, Nakasawa [1936b] observed that his axioms are equivalent to those of Whitney. The same axiom system as Nakasawa's, added with a continuity axiom, was given by Pauc [1937]. In Haupt, Nöbeling, and Pauc [1940] the concept of an *Abhängigkeitsraum* (dependence space) based on these axioms was investigated.

The fourth 'line' was that of Whitney [1935], who introduced the notion of a matroid as a concept by itself. He was motivated by generalizing certain separability and duality phenomena in graphs, studied by him before. This led him to show that each matroid has a dual. While Whitney showed the equivalence of several axiom systems for matroids, he did not consider an axiom system based on a closure operation or on flats. Whitney gave linear dependence as an example, but not algebraic dependence. In a paper in the same year and journal, Birkhoff [1935a] showed the relation of Whitney's work with lattices.

We now discuss some historical papers more extensively, in a more or less chronological order.

1894-1900: Dedekind: lattices

In the supplements to the fourth edition of *Vorlesungen über Zahlentheorie* (Lectures on Number Theory) by Lejeune Dirichlet [1894], R. Dedekind introduced the notion of a *module* as any nonempty set of (real or complex) numbers closed under addition and subtraction, and he studied the lattice of all modules ordered by inclusion. He called *A divisible by B* if $A \subseteq B$. Trivially, the lattice operations are given by $A \wedge B = A \cap B$ and $A \vee B = A + B$. In fact, Dedekind denoted $A \cap B$ by $A - B$.

He gave the following 'charakteristischen Satz' (characteristic theorem):

> *Ist m theilbar durch d, und a ein beliebiger Modul, so ist*
>
> $$m + (a - d) = (m + a) - d.$$ [3]

In modern notation, for all a, b, c:

(39.64) if $a \leq c$, then $a \vee (b \wedge c) = (a \vee b) \wedge c$,

which is now known as the *modular law*, and lattices obeying it are called *modular lattices*.

Next, Dedekind [1897] introduced the notion of a lattice under the name *Dualgruppe* (dual group), motivated by similarities observed by him between operations on modules and those for logical statements as given in the book *Algebra der Logik* (Algebra of Logic) by Schröder [1890]. Dedekind mentioned, as examples, subsets of a set, modules, ideals in a finite field, subgroups of a group, and all fields, and he introduced the name *module law* for property (39.64):

> ich will es daher das Modulgesetz nennen, und jede Dualgruppe, in welcher es herrscht, mag eine Dualgruppe vom Modultypus heißen.[4]

[3] *If m is divisible by d, and a is an arbitrary module, then*

$$m + (a - d) = (m + a) - d.$$

[4] I will therefore call it the module law, and every dual group in which it holds, may be called a dual group of module type.

Dedekind [1900] continued the study of modular lattices, and showed that each modular lattice allows a rank function $r : M \to \mathbb{Z}_+$ with the property that for all a, b:

(39.65) (i) $r(0) = 0$;
(ii) $r(b) = r(a) + 1$ if b covers a;
(iii) $r(a \wedge b) + r(a \vee b) = r(a) + r(b)$.

In fact, this characterizes modular lattices.

In proving (39.65), Dedekind showed that each modular lattice satisfies

(39.66) if a and b cover c, and $a \neq b$, then $a \vee b$ covers a and b,

which is the property characterizing *upper semimodular lattices*, a structure equivalent to matroids.

1862-1913: Grassmann, Steinitz: linear and algebraic dependence

The basic exchange property of linear independence was formulated by Grassmann [1862], in his book *Die Ausdehnungslehre*, as follows (in his terminology, vectors are quantities):

> **20.** *Wenn m Grössen $a_1, \ldots a_m$, die in keiner Zahlbeziehung zu einander stehen, aus n Grössen $b_1, \ldots b_n$ numerisch ableitbar sind, so kann man stets zu den m Grössen $a_1, \ldots a_m$ noch $(n-m)$ Grössen $a_{m+1}, \ldots a_n$ von der Art hinzufügen, dass sich die Grössen $b_1, \ldots b_n$ auch aus $a_1, \ldots a_n$ numerisch ableiten lassen, und also das Gebiet der Grössen $a_1, \ldots a_n$ identisch ist dem Gebiete der Grössen $b_1, \ldots b_n$; auch kann man jene $(n-m)$ Grössen aus den Grössen $b_1, \ldots b_n$ selbst entnehmen.*[5]

This property was also given by Steinitz [1913] (see below), but before that, Steinitz proved it for algebraic independence. In his fundamental paper *Algebraische Theorie der Körper* (Algebraic Theory of Fields), Steinitz [1910] studied, in § 22, algebraic dependence in field extensions. The statements proved are as follows, where L is a field extension of field K. Throughout, a is *algebraically dependent* on S if a is algebraic with respect to the field extension $K(S)$; in other words, if there is a nonzero polynomial $p(x) \in K(S)[x]$ with $p(a) = 0$.

Calling a set a *system*, he first observed:

> 1. *Hängt das Element a vom System S algebraisch ab, so gibt es ein endliches Teilsystem S' von S, von welchem a algebraisch abhängt.*[6]

and next he showed:

> 2. *Hängt S_3 von S_2, S_2 von S_1 algebraisch ab, so ist S_3 algebraisch abhängig von S_1.*[7]

---

[5] **20.** *If m quantities $a_1, \ldots a_m$, that stand in no number relation to each other, are numerically derivable from n quantities $b_1, \ldots b_n$, then one can always add to the m quantities $a_1, \ldots a_m$ another $(n-m)$ quantities $a_{m+1}, \ldots a_n$ such that the quantities $b_1, \ldots b_n$ can also be derived numerically from $a_1, \ldots a_n$, and that hence the domain of the quantities $a_1, \ldots a_n$ is identical to the domain of the quantities $b_1, \ldots b_n$; one also can take those $(n-m)$ quantities from the quantities $b_1, \ldots b_n$ themselves.*

[6] 1. *If element a depends algebraically on the system S, then there is a finite subsystem S' of S on which a depends algebraically.*

[7] 2. *If S_3 depends algebraically on S_2, and S_2 on S_1, then S_3 is algebraically dependent on S_1.*

He called two sets S_1 and S_2 *equivalent* if S_1 depends algebraically on S_2, and conversely. A set is *reducible* if it has a proper subset equivalent to it. He showed:

> 3. *Jedes Teilsystem eines irreduziblen Systems ist irreduzibel.*
> 4. *Jedes reduzible System enthält ein endliches reduzibles Teilsystem.*[8]

and (after statement 5, saying that any two field extensions by equicardinal irreducible systems are isomorphic):

> 6. *Wird ein irreduzibles System S durch Hinzufügung eines Elementes a reduzibel, so ist a von S algebraisch abhängig.*[9]

From these properties, Steinitz derived:

> 7. *Ist S ein (in bezug auf K) irreduzibles System, das Element a in bezug auf K transzendent, aber von S algebraisch abhängig, so enthält S ein bestimmtes endliches Teilsystem T von folgender Beschaffenheit: a ist von T algebraisch abhängig; jedes Teilsystem von S, von welchem a algebraisch abhängt, enthält das System T; wird irgendein Element aus T durch a ersetzt, so geht S in ein äquivalentes irreduzibles System über; keinem der übrigen Elemente von S kommt diese Eigenschaft zu.*[10]

Steinitz proved this using only the properties given above (together with the fact that any $s \in S$ is algebraically dependent on S). Moreover, he derived from 7, (what is now called) *Steinitz' exchange property* for algebraic dependence:

> 8. *Es seien U und B endliche irreduzible Systeme von m bzw. n Elementen; es sei $n \leq m$ und B algebraisch abhängig von U. Dann sind im Falle $m = n$ die Systeme U und B äquivalent, im Falle $n < m$ aber ist U einem irreduziblen System äquivalent, welches aus B und $m - n$ Elementen aus U besteht.*[11]

This in particular implies that any two equivalent irreducible systems have the same size, and that the properties are equivalent to that determining a matroid.

In a subsequent paper, Steinitz [1913] proved a number of auxiliary statements on linear equations. Among other things, he showed (in his terminology, vectors are numbers, and a vector space is a module):

> *Besitzt der Modul M eine Basis von p Zahlen, und enthält er r linear unabhängige Zahlen $\beta_1, \ldots, \beta_r$, so besitzt er auch eine Basis von p Zahlen, unter denen die Zahlen $\beta_1, \ldots, \beta_r$ sämtlich vorkommen.*[12]

---

[8] 3. *Every subsystem of an irreducible system is irreducible.*
4. *Every reducible system contains a finite reducible subsystem.*

[9] 6. *If an irreducible system S becomes reducible by adding an element a, then a is algebraically dependent on S.*

[10] 7. *If S is an irreducible system (with respect to K), [and] the element a transcendent with respect to K, but algebraically dependent on S, then S contains a certain finite subsystem T with the following quality: a is algebraically dependent on T; every subsystem of S on which a depends algebraically, contains the system T; if any element from T is replaced by a, then S passes into an equivalent irreducible system; this property belongs to none of the other elements of S .*

[11] 8. *Let U and B be finite irreducible systems of m and n elements respectively; let $n \leq m$ and let B be algebraically dependent on U. Then, in case $m = n$, the systems U and B are equivalent, but in case $n < m$, U is equivalent to an irreducible system which consists of B and $m - n$ elements from U.*

[12] *If a module M possesses a base of p numbers, and it contains r linearly independent numbers $\beta_1, \ldots, \beta_r$, then it possesses also a base of p numbers, among which the numbers $\beta_1, \ldots, \beta_r$ all occur.*

Steinitz' proof of this in fact gives a stronger result, known as *Steinitz' exchange property*: the new base is obtained by extending $\beta_1, \ldots, \beta_r$ with vectors from the given base. So Steinitz came to the same result as Grassmann [1862] quoted above. In his paper, Steinitz [1913] did not make a link with similar earlier results in Steinitz [1910] on algebraic dependence.

1929: Bergmann

Inspired by Menger [1928a], who aimed at giving an axiomatic foundation for projective geometry on a lattice-theoretical basis, Bergmann [1929] gave an axiomatic foundation of affine geometry, again on the basis of lattices. Bergmann's article contains a number of proofs that in fact concern matroids, while he assumed, but not used, a complementation axiom (since he aimed at characterizing full affine spaces, not subsets of it): for each pair of elements $A \leq B$ there exist C_1 and C_2 with $A \vee C_1 = B$, $A \wedge C_1 = 0$, $B \wedge C_2 = A$, and $B \vee C_2 = 1$. This obviously implies (in the finite case) that

(39.67) each element of the lattice is a join of points.

(A *point* is a minimal nonzero element.) It is property (39.67) that Bergmann uses in a number of subsequent arguments (and not the complementation axiom). His further axiom is:

(39.68) for any element A and any point P of the lattice, there is no element B with $A < B < A \vee P$.

He called an ordered sequence $(P_1, \ldots, P_n)$ of points a *chain* (Kette) (*of* an element A), if $P_i \not\leq P_1 \vee \cdots \vee P_{i-1}$ for $i = 1, \ldots, n$ (and $A = P_1 \vee \cdots \vee P_n$). He derived from (39.67) and (39.68) that being a chain is independent of the order of the elements in the chain, and that any two chains of an element A have the same length:

> *Satz: Alle Ketten eines Elementes A haben dieselbe Gliederzahl.*[13]

He remarked that under condition (39.67), this in turn implies (39.68).

Denoting the length of any chain of A by $|A|$, Bergmann showed that it is equal to the rank of A in the lattice, and he derived the submodular inequality:

$$|A| + |B| \geq |A + B| + |A \cdot B|.$$

(Bergmann denoted $\vee$ and $\wedge$ by $+$ and $\cdot$.) Thus he proved the submodularity of the rank function of a matroid. These results were also given by Alt [1936] in Menger's *mathematischen Kolloquium* in Vienna on 1 March 1935 (cf. Menger [1936a,1936b]).

1929-1937: Haupt, van der Waerden

Inspired by the work of Steinitz, in the books *Einführung in die Algebra* (Introduction to Algebra) by Haupt [1929a,1929b] and *Moderne Algebra* (Modern Algebra) by van der Waerden [1930], the analogies between proof methods for linear and algebraic dependence were observed.

Haupt mentioned in his preface (after saying that his book will contain the modern developments of algebra):

[13] *Theorem: All chains of an element A have the same number of members.*

> Demgemäß ist das vorliegende Buch durchweg beeinflußt von der bahnbrechenden ,,Algebraischen Theorie der Körper" von Herrn E. Steinitz, was hier ein für allemal hervorgehoben sei. Ferner stützt sich die Behandlung der linearen Gleichungen (vgl. *9,1* bis *9,4*), einer Anregung von Frl. E. Noether folgend, auf die von Herrn E. Steinitz gegebene Darstellung (vgl. das Zitat in *9,0*).[14]

(The quotation in Haupt's '*9,0*' is to Steinitz [1910,1913].)

A number of theorems on algebraic dependence were proved in Chapter 23 of Haupt [1929b] by referring to the proofs of the corresponding results on linear dependence in Chapter 9 of Haupt [1929a]. In the introduction of his Chapter 9, Haupt wrote:

> Die Behandlung der linearen Gleichungen ist (soweit es geht) so angelegt, daß sich ein Teil der dabei gewonnenen Sätze auf *Systeme von algebraisch abhängigen Elementen überträgt*, was später (*23,6*) dargelegt wird.[15]

In the first edition of his book, van der Waerden [1930] listed the properties of algebraic dependence:

> Die Relation der algebraischen Abhängigkeit hat demnach die folgenden Eigenschaften:
> 1. a ist abhängig von sich selbst, d.h. von der Menge $\{a\}$.
> 2. Ist a abhängig von M, so hängt es auch von jeder Obermenge von M ab.
> 3. Ist a abhängig von M, so ist a schon von einer endlichen Untermenge $\{m_1, \ldots, m_n\}$ von M (die auch leer sein kann) abhängig.
> 4. Wählt man diese Untermenge minimal, so ist jedes m_i von a und den übrigen m_j abhängig.
> Weiter gilt:
> 5. Ist a abhängig von M und jedes Element von M abhängig von N, so ist a abhängig von N.[16]

Following Steinitz, van der Waerden called two sets *equivalent* if each element of the one set depends algebraically on the other set, and vice versa, while a set is *irreducible* if no element of it depends algebraically on the remaining.

Using only the properties 1-5, van der Waerden derived that each set contains an irreducible set equivalent to it, and that if $M \subseteq N$, then each irreducible subset of M equivalent to M can be extended to an irreducible subset of N equivalent to N — in other words, inclusionwise minimal subsets of M equivalent to M are

[14] Accordingly, the present book is invariably influenced by the pioneering 'Algebraic Theory of Fields' by Mr E. Steinitz, which be emphasized here once and for all. Further, following a suggestion by Miss E. Noether, the treatment of linear equations (cf. *9,1* to *9,4*) leans on the presentation by Mr E. Steinitz (cf. the quotation in *9,0*).

[15] The treatment of linear equations is (as far as it goes) made such that a part of the theorems obtained therewith *transfers to systems of algebraically dependent elements*, which will be discussed later (*23,6*).

[16] The relation of algebraic dependence has therefore the following properties:
1. a is dependent on itself, that is, on the set $\{a\}$.
2. If a is dependent on M, then it also depends on every superset of M.
3. If a is dependent on M, then a is dependent already on a finite subset $\{m_1, \ldots, m_n\}$ of M (that can also be empty).
4. If one chooses this subset minimal, then every m_i is dependent on a and the remaining m_j.
Further it holds:
5. If a is dependent on M and every element of M is dependent on N, then a is dependent on N.

independent, and inclusionwise maximal independent subsets of M are equivalent to M.

Van der Waerden [1930] also showed that two equivalent irreducible systems have the same size, but in the proof he uses polynomials. This is not necessary, since the properties 1-5 determine a matroid.

Van der Waerden noticed the analogy with linear dependence, treated in his § 28, where he uses specific facts on linear equations:

> Tatsächlich gelten für die dort betrachtete lineare Abhängigkeit dieselben Regeln 1 bis 5, die für die algebraische Abhängigkeit in § 61 aufgestellt wurden; man kann also alle Beweise wörtlich übertragen.[17]

In the second edition of his book, van der Waerden [1937] gave a unified treatment of linear and algebraic dependence, slightly different from the first edition. As for linear dependence he stated in § 33:

> Drei Grundsätze genügen. Der erste ist ganz selbstverständlich.
> Grundsatz 1. *Jedes* u_i $(i = 1, \ldots, n)$ *ist von* $u_1, \ldots, u_n$ *linear abhängig.*
> Grundsatz 2. *Ist* v *linear abhängig von* $u_1, \ldots, u_n$, *aber nicht von* $u_1, \ldots, u_{n-1}$, *so ist* u_n *linear abhängig von* $u_1, \ldots, u_{n-1}, v$.
> [$\cdots$]
> Grundsatz 3. *Ist* w *linear abhängig von* $v_1, \ldots, v_s$ *und ist jedes* v_j $(j = 1, \ldots, s)$ *linear abhängig von* $u_1, \ldots, u_n$, *so ist* w *linear abhängig von* $u_1, \ldots, u_n$.[18]

The same axioms are given in § 64 of van der Waerden [1937], with 'linear' replaced by 'algebraisch'.

Next, van der Waerden called elements $u_1, \ldots, u_n$ (linearly or algebraically) independent if none of them depend on the rest of them. Among the consequences of these principles, he mentioned that if $u_1, \ldots, u_{n-1}$ are independent but $u_1, \ldots, u_{n-1}, u_n$ are not, then u_n is dependent on $u_1, \ldots, u_{n-1}$, and that each finite system of elements $u_1, \ldots, u_n$ contains a (possibly empty) independent subsystem on which each u_i is dependent. He called two systems $u_1, \ldots, u_n$ and $v_1, \ldots, v_s$ *equivalent* if each v_k depends on $u_1, \ldots, u_n$ and each u_i depends on $v_1, \ldots, v_s$, and he now derived from the three principles that two equivalent independent systems have the same size.

Mac Lane [1938] observed that the axioms introduced by Whitney [1935] and those by van der Waerden [1937] determine equivalent structures.

1934: Hilbert, Bernays: collinearity axioms

Axiom systems for points and lines in a plane were given by Hilbert [1899] in his book *Grundlagen der Geometrie* (Foundations of Geometry), and by Veblen [1904].

---

[17] In fact, the same rules 1 to 5, that were formulated for algebraic dependence in § 61, hold for the linear dependence considered there; one can transfer therefore all proofs word for word.

[18] Three principles suffice. The first one is fully self-evident.
Principle 1. *Every* u_i $(i = 1, \ldots, n)$ *is linearly dependent on* $u_1, \ldots, u_n$.
Principle 2. *If* v *is linearly dependent on* $u_1, \ldots, u_n$, *but not on* $u_1, \ldots, u_{n-1}$, *then* u_n *is linearly dependent on* $u_1, \ldots, u_{n-1}, v$.
[$\cdots$]
Principle 3. *If* w *is linearly dependent on* $v_1, \ldots, v_s$ *and every* v_j $(j = 1, \ldots, s)$ *is linearly dependent on* $u_1, \ldots, u_n$, *then* w *is linearly dependent on* $u_1, \ldots, u_n$.

Basis is the axiom that any two distinct points are in exactly one line. Note that this axiom determines precisely all matroids of rank at most 3 with no parallel elements (by taking the lines as maximal flats).

One of the axioms of Veblen is:

> Axiom VI. If points C and D ($C \neq D$) lie on the line AB, then A lies on the line CD.

This axiom corresponds to axiom 3) in the book *Grundlagen der Mathematik* (Foundations of Mathematics) of Hilbert and Bernays [1934], who aim to make an axiom system based on points only:

> Dabei empfiehlt es sich für unseren Zweck, von dem HILBERTschen Axiomensystem darin abzuweichen, daß wir nicht die Punkte und die Geraden als zwei Systeme von Dingen zugrunde legen, sondern *nur die Punkte als Individuen nehmen.*[19]

The axiom system of Hilbert and Bernays is in terms of a relation Gr to describe collinearity of triples of points (where (x) stands for $\forall x$, (Ex) for $\exists x$, and $\overline{P}$ for the negation of P):

I. Axiome der Verknüpfung.

1) $(x)(y)Gr(x,x,y)$
„x,x,y liegen stets auf einer Geraden."
2) $(x)(y)(z)(Gr(x,y,z) \rightarrow Gr(y,x,z) \& Gr(x,z,y))$.
„Wenn x,y,z auf einer Geraden liegen, so liegen stets auch y,x,z sowie auch x,z,y auf einer Geraden."
3) $(x)(y)(z)(u)(Gr(x,y,z) \& Gr(x,y,u) \& x \neq y \rightarrow Gr(x,z,u))$.
„Wenn x,y, verschiedene Punkte sind und wenn x,y,z sowie x,y,u auf einer Geraden liegen, so liegen stets auch x,z,u auf einer Geraden."
4) $(Ex)(Ey)(Ez)\overline{Gr(x,y,z)}$.
„Es gibt Punkte x,y,z, die nicht auf einer Geraden liegen."[20]

The axioms 1) and 2) in fact tell that the relation Gr is determined by unordered triples of distinct points. The exchange axiom 3) is a special case of the matroid axiom for circuits in a matroid.

Hilbert and Bernays extended the system by axioms for a betweenness relation Zw for ordered triples of points, and a parallelism relation Par for ordered quadruples of points.

[19] At that it is advisable for our purpose to deviate from HILBERT's axiom system in that we do not lay the points and the lines as two systems of things as base, but *take only the points as individuals.*

[20] I. Axioms of connection.

1) $(x)(y)Gr(x,x,y)$
'x,x,y always lie on a line.'
2) $(x)(y)(z)(Gr(x,y,z) \rightarrow Gr(y,x,z) \& Gr(x,z,y))$.
'If x,y,z lie on a line, then also y,x,z as well as x,z,y always lie on a line.'
3) $(x)(y)(z)(u)(Gr(x,y,z) \& Gr(x,y,u) \& x \neq y \rightarrow Gr(x,z,u))$.
'If x,y, are different points and if x,y,z as well as x,y,u lie on a line, then also x,z,u always lie on a line.'
4) $(Ex)(Ey)(Ez)\overline{Gr(x,y,z)}$.
'There are points x,y,z, that do not lie on a line.'

1933–1935: Birkhoff: Lattices

In his paper 'On the combination of subalgebras', Birkhoff [1933] ('Received 15 May 1933') wrote:

> The purpose of this paper is to provide a point of vantage from which to attack combinatorial problems in what may be termed modern, synthetic, or abstract algebra. In this spirit, a research has been made into the consequences and applications of seven or eight axioms, only one [V] of which itself is new.

The axioms are those for a lattice, added with axiom V, that amounts to (39.64) above. Any lattice satisfying this condition is called by Birkhoff in this paper a 'B-lattice'. In an addendum, Birkhoff [1934b] mentioned that O. Ore had informed him that part of his results had been obtained before by Dedekind [1900]. Therefore, Birkhoff [1935b] renamed it to *modular lattice*.

Birkhoff [1933] mentioned, as examples, the classes of normal subgroups and of characteristic subgroups of a group. Other examples mentioned are the ideals of a ring, and the linear subspaces of Euclidean space. (Both examples actually give sublattices of the lattice of all normal subgroups of the corresponding groups.)

Like Dedekind, Birkhoff [1933] showed that (39.64) implies (39.66). Lattices satisfying (39.66) are called (upper) *semimodular*. Birkhoff showed that any upper semimodular lattice has a rank function satisfying (39.65)(i) and (ii) and satisfying the submodular law:

(39.69) $\quad r(a \cap b) + r(a \cup b) \leq r(a) + r(b).$

This characterizes upper semimodular lattices.

Birkhoff noticed that this implies that the modular lattices are exactly those lattices satisfying both (39.66) and its symmetric form:

(39.70) $\quad$ if c covers a and b and $a \neq b$, then a and b cover $a \wedge b$.

Birkhoff [1935c] showed that the partition lattice is upper semimodular, that is, satisfies (39.66), and hence has a rank function satisfying the submodular inequality[21]. Thus the complete graph, and hence any graph, gives a geometric lattice (and hence a matroid — however, Whitney's work seems not to have been known yet to Birkhoff at the time of writing this paper).

In a number of other papers, Birkhoff [1934a,1934c,1935b] made a further study of modular lattices, and gave relations to projective geometries (in which the collection of all flats gives a modular lattice). Klein-Barmen [1937] further investigated semimodular lattices (called by him *Birkhoffsche Verbände* (Birkhoff lattices)), of which he found several lattice-theoretical characterizations.

1935: Whitney: Matroids

Whitney [1935] (presented to the American Mathematical Society, September 1934) introduces the notion of matroid as follows:

[21] In fact, Birkhoff [1935c] claimed the modular equality for the rank function of a partition lattice (page 448), but this must be a typo, witness the formulation of, and the reference in, the first footnote on that page.

> Let $C_1, C_2, \cdots, C_n$ be the columns of a matrix M. Any subset of these columns is either linearly independent or linearly dependent; the subsets thus fall into two classes. These classes are not arbitrary; for instance, the two following theorems must hold:
>
> (a) Any subset of an independent set is independent.
>
> (b) If N_p and N_{p+1} are independent sets of p and $p+1$ columns respectively, then N_p together with some column of N_{p+1} forms an independent set of $p+1$ columns.
>
> There are other theorems not deducible from this; for in § 16 we give an example of a system satisfying these two theorems but not representing any matrix. Further theorems seem, however, to be quite difficult to find. Let us call a system obeying (a) and (b) a "matroid." The present paper is devoted to a study of the elementary properties of matroids. The fundamental question of completely characterizing systems which represent matrices is left unsolved. In place of the columns of a matrix we may equally well consider points or vectors in a Euclidean space, or polynomials, etc.

In the paper, Whitney observed that forests in a graph form the independent sets of a matroid, for which reason he carried over various terms from graphs to matroids.

Whitney described several equivalent axiom systems for the notion of matroid. First, he showed that the rank function is characterized by (39.42), and he derived that it is submodular. Next, he showed that the collection of bases is characterized by (39.33)(ii), and the collection of circuits by (39.34)(iii). Moreover, he showed that complementing all bases gives again a matroid, the dual matroid, and that the dual of a linear matroid is again a linear matroid. In the paper, he also studied separability and representability of matroids. The example given in Whitney's § 16 (mentioned in the above quotation), is in fact the well-known Fano matroid — he apparently did not consider matrices over GF(2). However, in an appendix of the paper, he characterized the matroids representable by a matrix 'of integers mod 2': a matroid is representable over GF(2) if and only if any sum (mod 2) of circuits can be partitioned into circuits.

In a subsequent paper 'Abstract linear independence and lattices', Birkhoff [1935a] pointed out the relations of Whitney's work with Birkhoff's earlier work on semimodular lattices. He stated:

> In a preceding paper, Hassler Whitney has shown that it is difficult to distinguish theoretically between the properties of linear dependence of ordinary vectors, and those of elements of a considerably wider class of systems, which he has called "matroids."
>
> Now it is obviously impossible to incorporate all of the heterogeneous abstract systems which are constantly being invented, into a body of systematic theory, until they have been classified into two or three main species. The purpose of this note is to correlate matroids with abstract systems of a very common type, which I have called "lattices."

Birkhoff showed that a lattice is isomorphic to the lattice of flats of a matroid if and only if the lattice is semimodular, that is, satisfies (39.66), and each element is a join of atoms.

In the paper 'Some interpretations of abstract linear dependence in terms of projective geometry', MacLane [1936] gave a geometric interpretation of matroids. He introduced the notion of a 'schematic n-dimensional figure', consisting of 'k-dimensional planes' for $k = 1, 2, \ldots$. Each such plane is a subset of an (abstract) set of 'points', with the following axioms (for any appropriate k):

(39.71) (i) any k points belonging to no $k-1$-dimensional plane, belong to a unique k-dimensional plane; moreover, this plane is contained in any plane containing these k points;
(ii) every k-dimensional plane contains k points that belong to no $k-1$-dimensional plane.

MacLane mentioned that there is a 1-1 correspondence between schematic figures and the collections of flats of matroids. As a consequence he mentioned that a schematic n-dimensional figure is completely determined by its collection of $n-1$-dimensional planes (as a matroid is determined by its hyperplanes = complements of cocircuits).

1935: Nakasawa: Abhängigkeitsräume

In the paper *Zur Axiomatik der linearen Abhängigkeit. I* (On the axiomatics of linear dependence. I) in *Science Reports of the Tokyo Bunrika Daigaku* (Tokyo University of Literature and Science), Nakasawa [1935] introduced an axiom system for dependence, that he proved to be equivalent to matroids (in a different terminology).

He was motivated by an axiom system described by Thomsen [1933] in his book *Grundlagen der Elementargeometrie* (Foundations of Elementary Geometry). Thomsen's 'cycle calculus' is an attempt to axiomatize relations (like coincidence, orthogonality, parallelism) between geometric objects (points, lines, etc.). Thomsen emphasized that existence questions often are inessential in elementary geometry:

> In der Tat erscheinen uns ja auch die Existenzaussagen als ein verhältnismäßig unwesentliches Beiwerk der Elementargeometrie. Ohne Zweifel empfinden wir als die eigentlich inhaltsvollsten und die wichtigsten Einzelaussagen der Elementargeometrie die von der folgenden reinen Form: ,,Wenn eine Reihe von geometrischen Gebilden, d.h. eine Anzahl von Punkten, Geraden, usw., gegeben vorliegt, und zwar derart, daß zwischen den gegebenen Punkten, Geraden usw. die und die geometrischen Lagebeziehungen bestehen (Koinzidenz, Senkrechtstehen, Parallellaufen, ,,Mittelpunkt sein" und anderes mehr), dann ist eine notwendige Folge dieser Annahme, daß auch noch diese bestimmte weitere geometrische Lagebeziehung gleichzeitig besteht." In Sätzen dieser Form kommt nichts von Existenzaussagen vor. Was das Wichtigste ist, nicht in den Folgerungen. Dann aber auch nicht in den Annahmen. Wir nehmen an: *Wenn* die und die Dinge in den und den Beziehungen gegeben vorliegen..., usw. Wir machen aber keinerlei Voraussetzungen darüber, ob eine solche Konfiguration in unserer Geometrie existieren kann. Der Schluß ist nur: Wenn sie existieren, dann .... Falls die Konfiguration gar nicht existiert, der Satz also gegenstandslos wird, betrachten wir ihn nach der üblichen Konvention ,,gegenstandslos, also richtig" als richtig.[22]

[22] Indeed, also the existence statements seem to us a relatively inessential side issue of elementary geometry. Undoubtedly, we find as the really most substantial and most important special statements of elementary geometry those of the following pure form: 'If a sequence of geometric creations, that is, a number of points, lines etc., are given to us, and that in such a way, that those and those geometric position relations exist between the given points, lines etc. (coincidence, orthogonality, parallelism, "being a centre", and other), then a necessary consequence of this assumption is that also this certain further geometric position relation exists at the same time.' In theorems of this form, no existence statements occur. What is most important: not in the consequences. But then neither in the assumptions. We assume: *If* those and those things are given

Thomsen aimed at founding axiomatically 'the partial geometry of all elementary geometric theorems without existence statements'. To that end, he introduced the concept of a *cycle*, which is an ordered finite sequence of abstract objects, which can be thought of as points, lines, etc. Certain cycles are 'correct' and the other 'incorrect' (essentially they represent a system of relations defining any binary group):

> A) *Axiom der Grundzyklen:* Der Zyklus $\alpha\alpha$ ist für jedes α richtig, der Zyklus α für kein α.
> B) *Axiom des Löschens:* $\beta_1\beta_2\ldots\beta_n\alpha\alpha \rightarrow \beta_1\beta_2\ldots\beta_n$; in Worten: Aus der Richtigkeit des Zyklus $\beta_1\beta_2\ldots\beta_n\alpha\alpha$ folgt auch die des Zyklus $\beta_1\beta_2\ldots\beta_n$.
> C) *Axiom des Umstellens:* $\beta_1\beta_2\ldots\beta_n \rightarrow \beta_2\beta_3\ldots\beta_n\beta_1$.
> D) *Axiom des Umkehrens:* $\beta_1\beta_2\ldots\beta_{n-1}\beta_n \rightarrow \beta_n\beta_{n-1}\ldots\beta_2\beta_1$.
> E) *Axiom des Anfügens:* $\beta_1\beta_2\ldots\beta_n$ und $\gamma_1\gamma_2\ldots\gamma_r \rightarrow \beta_1\beta_2\ldots\beta_n\gamma_1\gamma_2\ldots\gamma_r$.[23]

Axiom B) can be considered as a variant of Steinitz' exchange property. With the other axioms it implies that if $\beta_1\cdots\beta_n\alpha$ and $\gamma_1\cdots\gamma_r\alpha$ are cycles, then $\beta_1\cdots\beta_n\gamma_1\cdots\gamma_r$ is a cycle. Therefore, the set of all inclusionwise minimal nonempty sets containing a cycle form the circuits of a matroid.

The purpose of Nakasawa [1935] is to generalize Thomsen's axiom system:

> In der vorliegenden Untersuchung soll ein Axiomensystem für eine neue Formulierung der linearen Abhängigkeit des n-dimensionalen projektiven Raumes angegeben werden, indem wir hauptsächlich den *Zyklenkalkül*, den Herr G. Thomsen bei seiner Grundlegung der elementaren Geometrie hergestellt hat, hier in einem noch abstrakteren Sinne verwenden.[24]

While Thomsen's cycles relate to unions of circuits in a matroid, those of Nakasawa form the dependent sets of a matroid. His axiom system can be considered as a direct extension to higher dimensions of the collinearity axioms of Hilbert and Bernays given above.

He called the structure *der erste Verknüpfungsraum* (the first connection space), or a $\mathcal{B}_1$-*Raum* ($\mathcal{B}_1$-space), writing $a_1\cdots a_s$ for $a_1\cdots a_s = 0$:

> **Grundannahme:** Wir denken uns eine gewisse Menge der Elementen; $\mathcal{B}_1 \ni a_1, a_2, \cdots, a_s, \cdots$. Für gewisse Reihen der Elementen, die wir *Zyklen* nennen wollen, denken wir dazu die Relationen "*gelten*" oder "*gültig sein*", in Zeichen $a_1\cdots a_s = 0$, bzw. "*nicht gelten*" oder "*nicht gültig sein*", in Zeichen $a_1\cdots a_s \neq 0$. Diese Relationen sollen nun folgenden Axiomen genügen;

to us in those and those relations..., etc. We do not make any assumption on the fact if such a configuration can exist in our geometry. The conclusion is only: If they exist, then .... In case the configuration does not exist at all, and the theorem thus becomes meaningless, we consider it by the usual convention 'meaningless, hence correct' as correct.

[23]

A) *Axiom of ground cycles:* The cycle $\alpha\alpha$ is correct for each α, the cycle α for no α.
B) *Axiom of solving:* $\beta_1\beta_2\ldots\beta_n\alpha\alpha \rightarrow \beta_1\beta_2\ldots\beta_n$; in words: From the correctness of the cycle $\beta_1\beta_2\ldots\beta_n\alpha\alpha$ follows that of the cycle $\beta_1\beta_2\ldots\beta_n$.
C) *Axiom of transposition:* $\beta_1\beta_2\ldots\beta_n \rightarrow \beta_2\beta_3\ldots\beta_n\beta_1$.
D) *Axiom of inversion:* $\beta_1\beta_2\ldots\beta_{n-1}\beta_n \rightarrow \beta_n\beta_{n-1}\ldots\beta_2\beta_1$.
E) *Axiom of addition:* $\beta_1\beta_2\ldots\beta_n$ and $\gamma_1\gamma_2\ldots\gamma_r \rightarrow \beta_1\beta_2\ldots\beta_n\gamma_1\gamma_2\ldots\gamma_r$.

[24] In the present research, an axiom system for a new formulation of linear dependence of the n-dimensional projective space should be indicated, while we use here mainly the *cycle calculus*, which Mr G. Thomsen has constructed in his foundation of elementary geometry, in a still more abstract sense.

Axiom 1. (Reflexivität) : aa.
Axiom 2. (Folgerung) : $a_1 \cdots a_s \to a_1 \cdots a_s x$, $(s = 1, 2, \cdots)$.
Axiom 3. (Vertauschung) : $a_1 \cdots a_i \cdots a_s \to a_i \cdots a_1 \cdots a_s$, $(s = 2, 3, \cdots; i = 2, \cdots, s)$.
Axiom 4. (Transitivität) : $a_1 \cdots a_s \neq 0$, $xa_1 \cdots a_s$, $a_1 \cdots a_s y$ $\to xa_1 \cdots a_{s-1} y$, $(s = 1, 2, \cdots)$.

Definition I. Eine solche Menge $\mathcal{B}_1$ heisst der erste Verknüpfungsraum, in kurzen Worten, $\mathcal{B}_1$-**Raum**.[25]

Axiom 3 corresponds to condition (39.3).

Nakasawa introduced the concept of span, and he derived that any two independent sets having the same span, have the same size. It implies that $\mathcal{B}_1$-spaces are the same structures as matroids. Moreover, he gave a submodular law for a rank concept.

In a second paper, Nakasawa [1936a] added a further axiom on intersections of subspaces, yielding a '$\mathcal{B}_2$-space', which corresponds to a projective space (in which the rank is modular), and in a third paper, Nakasawa [1936b] observed that his $\mathcal{B}_1$-spaces form the same structure as the matroids of Whitney.

1937-1940: Pauc, Haupt, Nöbeling

The axioms presented by Nakasawa were also given by Pauc [1937], added with an axiom describing the limit behaviour of dependence, if the underlying set is endowed with a topology:

INTRODUCTION AXIOMATIQUE D'UNE NOTION DE DÉPENDANCE SUR UNE CLASSE LIMITE. — Soit D un prédicat relatif aux systèmes finis non ordonnés de points d'une classe limite $\mathcal{L}$, assujetti aux axiomes (notation d'Hilbert-Bernays)

(A$_1$) $(x_1)(x_2)(D[x_1, x_2] \sim (x_1 = x_2))$,

(A$_2$) $(x_1)(x_2)\ldots(x_p)(y)(D[x_1, x_2, \ldots, x_p] \to D[x_1, x_2, \ldots, x_p, y])$,

(A$_3$) $(x_1)(x_2)\ldots(x_p)(y)(z)(\overline{D[x_1, \ldots, x_p]}\&\ D[x_1, \ldots, x_p, y]\&$ $D[x_1, \ldots, x_p, z] \to D[x_2, \ldots, x_p, y, z])$,

(A$_4$) $\begin{cases} \text{Quels que soient les points } x_1, x_2, \ldots, x_p \text{ et la suite } y_1, y_2, \ldots, y_q, \\ \ldots \text{ de } \mathcal{L} \\ (\lim_{q\to\infty} y_q = y)\&(q)D[x_1, x_2, \ldots, x_p, y_q] \to D[x_1, x_2, \ldots, x_p, y]. \end{cases}$[26]

In a subsequent paper, Haupt, Nöbeling, and Pauc [1940] studied systems, called *A-Mannigfaltigkeit*, (*A*-manifolds) that satisfy the axioms A$_1$-A$_3$. They mentioned

[25] **Basic assumption:** We imagine ourselves a certain set of elements; $\mathcal{B}_1 \ni a_1, a_2, \cdots, a_s, \cdots$. For certain sequences of the elements, which we want to call *cycles*, we think the relations on them '*to hold*' or '*to be valid*', in notation $a_1 \cdots a_s = 0$, and '*not to hold*' or '*not to be valid*', in notation $a_1 \cdots a_s \neq 0$, respectively. These relations now should satisfy the following axioms;

Axiom 1. (reflexivity) : aa.
Axiom 2. (deduction) : $a_1 \cdots a_s \to a_1 \cdots a_s x$, $(s = 1, 2, \cdots)$.
Axiom 3. (exchange) : $a_1 \cdots a_i \cdots a_s \to a_i \cdots a_1 \cdots a_s$, $(s = 2, 3, \cdots; i = 2, \cdots, s)$.
Axiom 4. (transitivity) : $a_1 \cdots a_s \neq 0$, $xa_1 \cdots a_s$, $a_1 \cdots a_s y$ $\to xa_1 \cdots a_{s-1} y$, $(s = 1, 2, \cdots)$.

Definition I. Such a set $\mathcal{B}_1$ is called the first connection space, in short, $\mathcal{B}_1$-**space**.

[26] AXIOMATIC INTRODUCTION OF A NOTION OF DEPENDENCE ON A LIMIT CLASS. — Let D be a predicate relative to the finite unordered systems of points from a limit class $\mathcal{L}$, subject to the axioms (notation of Hilbert-Bernays)

that this axiom system was indeed inspired by those for collinearity of Hilbert-Bernays quoted above. They commented that its relation with Birkhoff's lattices, is analogous to the relation of the Hilbert-Bernays collinearity axioms with those of Hilbert for points and lines.

Haupt, Nöbeling, and Pauc [1940] gave, as examples, linear and algebraic dependence, and derived several basic facts (all bases have the same size, each independent set is contained in a base, for each pair of bases B, B' and $x \in B \setminus B'$ there is a $y \in B' \setminus B$ such that $B - x + y$ is a base, and the rank is submodular).

The authors mentioned that they were informed by G. Köthe about the relations of their work with the lattice formulation of algebraic dependence of Mac Lane [1938], but no connection is made with Whitney's matroid.

Among the further papers related to matroids are Menger [1936b], giving axioms for (full) affine spaces, and Wilcox [1939,1941,1942,1944] and Dilworth [1941a, 1941b,1944] on matroid lattices. The notion of M-symmetric lattice introduced by Wilcox [1942] was shown in Wilcox [1944] to be equivalent to upper semimodular lattice.

Rado

Rado was one of the first to take the independence structure as a source for further theorems, and to connect it with matching type theorems and combinatorial optimization. He had been interested in Kőnig-Hall type theorems (Rado [1933,1938]), and in his paper Rado [1942], he extended Hall's marriage theorem to transversals that are independent in a given matroid — a precursor of matroid intersection. In fact, with an elementary construction, Rado's theorem implies the matroid union theorem, and hence also the matroid intersection theorem (to be discussed in Chapters 41 and 42).

Rado [1942] did not refer to any earlier literature when introducing the concept of an *independence relation*, but the axioms are similar to those of Whitney for the independent sets in a matroid. Rado mentioned only linear independence as a special case.

He proved that a family of subsets of a matroid has an independent transversal if and only if the union of any k of the subsets contains an independent set of size k, for all k. Rado also showed that this theorem characterizes matroids.

Rado [1949a] extended the concept of matroid to infinite matroids, where he says that he extends the axioms of Whitney [1935].

Rado [1957] showed that if the elements of a matroid are linearly ordered by $\leq$, there is a unique minimal base $\{b_1, \dots, b_r\}$ with $b_1 < b_2 < \cdots < b_r$ such that for each $i = 1, \dots, r$ all elements $s < b_i$ belong to $\mathrm{span}(\{b_1, \dots, b_{i-1}\})$. Rado derived that for any independent set $\{a_1, \dots, a_k\}$ with $a_1 < \cdots < a_k$ one has $b_i \leq a_i$ for $i = 1, \dots, k$. Therefore, the greedy method gives an optimum solution when

---

(A1) $(x_1)(x_2)(D[x_1, x_2] \sim (x_1 = x_2))$,

(A2) $(x_1)(x_2)\dots(x_p)(y)(D[x_1, x_2, \dots, x_p] \to D[x_1, x_2, \dots, x_p, y])$,

(A3) $(x_1)(x_2)\dots(x_p)(y)(z)(\overline{D[x_1, \dots, x_p]}\&\ D[x_1, \dots, x_p, y]\&\ D[x_1, \dots, x_p, z] \to D[x_2, \dots, x_p, y, z])$,

(A4) $\begin{cases} \text{Whatever are the points } x_1, x_2, \dots, x_p \text{ and the sequence } y_1, y_2, \dots, y_q, \\ \dots \text{ from } \mathcal{L} \\ (\lim_{q \to \infty} y_q = y)\&(q)D[x_1, x_2, \dots, x_p, y_q] \to D[x_1, x_2, \dots, x_p, y]. \end{cases}$

applied to find a minimum-weight base. Rado mentioned that it extends the work of Borůvka and Kruskal on finding a shortest spanning tree in a graph.

For notes on the history of matroid union, see Section 42.6f. For an excellent survey of early literature on matroids, with reprints of basic articles, see Kung [1986].

Chapter 40

The greedy algorithm and the independent set polytope

We now pass to algorithmic and polyhedral aspects of matroids. We show that the greedy algorithm characterizes matroids and that it implies a characterization of the independent set polytope (the convex hull of the incidence vectors of the independent sets).
Algorithmic and polyhedral aspects of the *intersection* of two matroids will be studied in Chapter 41.

40.1. The greedy algorithm

Let $\mathcal{I}$ be a nonempty collection of subsets of a finite set S closed under taking subsets. For any weight function $w : S \to \mathbb{R}$ we want to find a set I in $\mathcal{I}$ maximizing $w(I)$. The *greedy algorithm* consists of setting $I := \emptyset$, and next repeatedly choosing $y \in S \setminus I$ with $I \cup \{y\} \in \mathcal{I}$ and with $w(y)$ as large as possible. We stop if no such y exists.

For general collections $\mathcal{I}$ of this kind this need not lead to an optimum solution. Indeed, matroids are precisely the structures where it always works, as the following theorem shows (Rado [1957] (necessity) and Gale [1968] and Edmonds [1971] (sufficiency)):

Theorem 40.1. *Let $\mathcal{I}$ be a nonempty collection of subsets of a set S, closed under taking subsets. Then the pair $(S, \mathcal{I})$ is a matroid if and only if for each weight function $w : S \to \mathbb{R}_+$, the greedy algorithm leads to a set I in $\mathcal{I}$ of maximum weight $w(I)$.*

Proof. *Necessity.* Let $(S, \mathcal{I})$ be a matroid and let $w : S \to \mathbb{R}_+$ be any weight function on S. Call an independent set I *good* if it is contained in a maximum-weight base. It suffices to show that if I is good, and y is an element in $S \setminus I$ with $I + y \in \mathcal{I}$ and with $w(y)$ as large as possible, then $I + y$ is good.

As I is good, there exists a maximum-weight base $B \supseteq I$. If $y \in B$, then $I + y$ is good again. If $y \notin B$, then there exists a base B' containing $I + y$ and contained in $B + y$. So $B' = B - z + y$ for some $z \in B \setminus I$. As $w(y)$ is chosen maximum and as $I + z \in \mathcal{I}$ since $I + z \subseteq B$, we know $w(y) \geq w(z)$.

Hence $w(B') \geq w(B)$, and therefore B' is a maximum-weight base. So $I + y$ is good.

Sufficiency. Suppose that the greedy algorithm leads to an independent set of maximum weight for each weight function $w : S \to \mathbb{R}_+$. We show that $(S, \mathcal{I})$ is a matroid.

Condition (39.1)(i) is satisfied by assumption. To see condition (39.1)(ii), let $I, J \in \mathcal{I}$ with $|I| < |J|$. Suppose that $I + z \notin \mathcal{I}$ for each $z \in J \setminus I$.

Let $k := |I|$. Consider the following weight function w on S:

$$w(s) := \begin{cases} k+2 & \text{if } s \in I, \\ k+1 & \text{if } s \in J \setminus I, \\ 0 & \text{if } s \in S \setminus (I \cup J). \end{cases} \tag{40.1}$$

Now in the first k iterations of the greedy algorithm we find the k elements in I. By assumption, at any further iteration, we cannot choose any element in $J \setminus I$. Hence any further element chosen, has weight 0. So the greedy algorithm yields an independent set of weight $k(k+2)$.

However, J has weight at least $|J|(k+1) \geq (k+1)(k+1) > k(k+2)$. Hence the greedy algorithm does not give a maximum-weight independent set, contradicting our assumption. ∎

The theorem restricts w to nonnegative weight functions. However, it is shown similarly that for matroids $M = (S, \mathcal{I})$ and arbitrary weight functions $w : S \to \mathbb{R}$, the greedy algorithm finds a maximum-weight base. By replacing 'as large as possible' in the greedy algorithm by 'as small as possible', one obtains an algorithm finding a *minimum*-weight base in a matroid. Moreover, by deleting elements of negative weight, the algorithm can be adapted to yield an independent set of maximum weight, for any weight function $w : S \to \mathbb{R}$.

Throughout we assume that the matroid $M = (S, \mathcal{I})$ is given by an algorithm testing if a given subset of S belongs to $\mathcal{I}$. We call this an *independence testing oracle*. So the full list of all independent sets is not given explicitly (such a list would increase the size of the input exponentially, making most complexity issues meaningless).

In explicit applications, the matroid usually can be described by such a polynomial-time algorithm (polynomial in $|S|$). For instance, we can test if a given set of edges of a graph $G = (V, E)$ is a forest in time polynomially bounded by $|V| + |E|$. So the matroid $(E, \mathcal{F})$ can be described by such an algorithm.

Under these assumptions we have:

Corollary 40.1a. *A maximum-weight independent set in a matroid can be found in strongly polynomial time.*

Proof. See above. ∎

Similarly, for minimum-weight bases:

Corollary 40.1b. *A minimum-weight base in a matroid can be found in strongly polynomial time.*

Proof. See above. ∎

40.2. The independent set polytope

The algorithmic results obtained in the previous section have interesting consequences for polyhedra associated with matroids, as was shown by Edmonds [1970b,1971,1979].

The *independent set polytope* $P_{\text{independent set}}(M)$ of a matroid $M = (S, \mathcal{I})$ is, by definition, the convex hull of the incidence vectors of the independent sets of M. So $P_{\text{independent set}}(M)$ is a polytope in $\mathbb{R}^S$.

Each vector x in $P_{\text{independent set}}(M)$ satisfies the following linear inequalities:

(40.2)
$$\begin{aligned} x_s &\geq 0 && \text{for } s \in S, \\ x(U) &\leq r_M(U) && \text{for } U \subseteq S, \end{aligned}$$

because the incidence vector χ^I of any independent set I of M satisfies (40.2). Note that x is an integer vector satisfying (40.2) if and only if x is the incidence vector of some independent set of M.

Edmonds showed that system (40.2) fully determines the independent set polytope, by deriving it from the following formula (yielding a good characterization):

Theorem 40.2. *Let $M = (S, \mathcal{I})$ be a matroid, with rank function r. Then for any weight function $w : S \to \mathbb{R}_+$:*

(40.3)
$$\max\{w(I) \mid I \in \mathcal{I}\} = \sum_{i=1}^{n} \lambda_i r(U_i),$$

where $U_1 \subset \cdots \subset U_n \subseteq S$ and where $\lambda_i \geq 0$ satisfy

(40.4)
$$w = \sum_{i=1}^{n} \lambda_i \chi^{U_i}.$$

Proof. Order the elements of S as $s_1, \ldots, s_n$ such that $w(s_1) \geq w(s_2) \geq \cdots \geq w(s_n)$. Define

(40.5)
$$U_i := \{s_1, \ldots, s_i\}$$

for $i = 0, \ldots, n$, and

(40.6)
$$I := \{s_i \mid r(U_i) > r(U_{i-1})\}.$$

So I is the output of the greedy algorithm. Hence I is a maximum-weight independent set.

Next let:

(40.7) $$\lambda_i := w(s_i) - w(s_{i+1}) \text{ for } i = 1, \ldots, n-1, \\ \lambda_n := w(s_n).$$

This implies (40.3):

(40.8) $$w(I) = \sum_{s \in I} w(s) = \sum_{i=1}^{n} w(s_i)(r(U_i) - r(U_{i-1})) \\ = w(s_n)r(U_n) + \sum_{i=1}^{n-1}(w(s_i) - w(s_{i+1}))r(U_i) = \sum_{i=1}^{n} \lambda_i r(U_i).$$

By taking any ordering of S for which w is nonincreasing, (40.5) gives any chain of subsets U_i satisfying (40.4) for some $\lambda_i \geq 0$. Hence we have the theorem. ∎

This can be interpreted in terms of LP-duality. For any weight function $w : S \to \mathbb{R}$, consider the linear programming problem

(40.9) $$\begin{array}{lll} \text{maximize} & w^{\mathsf{T}}x, & \\ \text{subject to} & x_s \geq 0 & (s \in S), \\ & x(U) \leq r_M(U) & (U \subseteq S), \end{array}$$

and its dual:

(40.10) $$\begin{array}{lll} \text{minimize} & \displaystyle\sum_{U \subseteq S} y_U r_M(U), & \\ \text{subject to} & y_U \geq 0 & (U \subseteq S), \\ & \displaystyle\sum_{U \subseteq S} y_U \chi^U \geq w. & \end{array}$$

Corollary 40.2a. *If $w : S \to \mathbb{Z}$, then* (40.9) *and* (40.10) *have integer optimum solutions.*

Proof. We can assume that $w(s) \geq 0$ for each $s \in S$ (as neither the maximum nor the minimum changes by resetting $w(s)$ to 0 if negative). Then (40.4) implies that the λ_i are integer. This gives integer optimum solutions of (40.9) and (40.10). ∎

In polyhedral terms, Theorem 40.2 implies:

Corollary 40.2b. *The independent set polytope is determined by* (40.2).

Proof. Immediately from Theorem 40.2 (with (40.10)). ∎

Moreover, in TDI terms:

Corollary 40.2c. *System* (40.2) *is totally dual integral.*

Proof. Immediately from Corollary 40.2a. ∎

Similar results hold for the base polytope. For any matroid M, let $P_{\text{base}}(M)$ be the *base polytope* of M, defined as the convex hull of the incidence vectors of bases of M. Then:

Corollary 40.2d. *The base polytope of a matroid* $M = (S, \mathcal{I})$ *is determined by*

(40.11)
$$\begin{array}{ll} x_s \geq 0 & \text{for } s \in S, \\ x(U) \leq r_M(U) & \text{for } U \subseteq S, \\ x(S) = r_M(S). & \end{array}$$

Proof. This follows directly from Corollary 40.2b, since the base polytope is the intersection of the independent set polytope with the hyperplane $\{x \mid x(S) = r_M(S)\}$, as an independent set I is a base if and only if $|I| \geq r_M(S)$. ∎

The corresponding TDI result reads:

Corollary 40.2e. *System* (40.11) *is totally dual integral.*

Proof. By Theorem 5.25 from Corollary 40.2c. ∎

One can similarly describe the *spanning set polytope* $P_{\text{spanning set}}(M)$ of M, which is, by definition, the convex hull of the incidence vectors of the spanning sets of M. It is determined by the system:

(40.12)
$$\begin{array}{ll} 0 \leq x_s \leq 1 & \text{for } s \in S, \\ x(U) \geq r_M(S) - r_M(S \setminus U) & \text{for } U \subseteq S. \end{array}$$

Corollary 40.2f. *The spanning set polytope is determined by* (40.12).

Proof. A subset U of S is spanning in M if and only if $S \setminus U$ is independent in M^*. Hence for any $x \in \mathbb{R}^S$ we have:

(40.13)
$$x \in P_{\text{spanning set}}(M) \iff \mathbf{1} - x \in P_{\text{independent set}}(M^*).$$

By Corollary 40.2b, $\mathbf{1} - x$ belongs to $P_{\text{independent set}}(M^*)$ if and only if x satisfies:

(40.14)
$$\begin{array}{ll} 1 - x_s \geq 0 & \text{for } s \in S, \\ |U| - x(U) \leq r_{M^*}(U) & \text{for } U \subseteq S. \end{array}$$

Since $r_{M^*}(U) = |U| + r_M(S \setminus U) - r_M(S)$, the present corollary follows. ∎

Corollary 40.2c gives similarly the TDI result:

Corollary 40.2g. *System* (40.12) *is totally dual integral.*

Proof. By reduction to Corollary 40.2c, by a similar reduction as in the proof of the previous corollary. ∎

Note that

(40.15) $$P_{\text{base}}(M) = P_{\text{independent set}}(M) \cap P_{\text{spanning set}}(M),$$
$$P_{\text{independent set}}(M) = P^{\downarrow}_{\text{base}}(M) \cap [0,1]^S,$$
$$P_{\text{spanning set}}(M) = P^{\uparrow}_{\text{base}}(M) \cap [0,1]^S.$$

The following consequence on the intersection of the base polytope with a box was observed by Hell and Speer [1984]:

Corollary 40.2h. *Let $M = (S, \mathcal{I})$ be a matroid and let $l, u \in \mathbb{R}^S$ with $l \leq u$. Then there is an $x \in P_{\text{base}}(M)$ with $l \leq x \leq u$ if and only if $l \in P^{\downarrow}_{\text{base}}(M)$ and $u \in P^{\uparrow}_{\text{base}}(M)$.*

Proof. Necessity being trivial, we show sufficiency. We may assume that $l, u \in [0,1]^S$. So $l \in P_{\text{independent set}}(M)$ and $u \in P_{\text{spanning set}}(M)$. Choose l', u' such that $l \leq l' \leq u' \leq u$, $l' \in P_{\text{independent set}}(M)$, $u' \in P_{\text{spanning set}}(M)$, and $\|u' - l'\|_1$ minimal.

If $l' = u'$ we are done, so assume that there is an $s \in S$ with $l'(s) < u'(s)$. As we cannot increase $l'(s)$, there is a $T \subseteq S$ with $s \in T$ and $l'(T) = r(T)$. Similarly, as we cannot decrease $u'(s)$, there is a $U \subseteq S$ with $s \notin U$ and $u'(S \setminus U) = r(S) - r(U)$. Then we have the contradiction

(40.16) $$\begin{aligned} &l'(T \cap U) + u'(T \cup U) \leq r(T \cap U) + u'(S) + r(T \cup U) - r(S) \\ &\leq r(T) + r(U) + u'(S) - r(S) = l'(T) + u'(U) \\ &< l'(T \cap U) + u'(T \cup U). \end{aligned}$$

The last inequality follows from

(40.17) $$u'(T \cup U) - u'(U) = u'(T \setminus U) > l'(T \setminus U) = l'(T) - l'(T \cap U),$$

since $s \in T \setminus U$ and $u'(s) > l'(s)$. ∎

40.3. The most violated inequality

We now consider the problem to find, for any matroid $M = (S, \mathcal{I})$ and any $x \in \mathbb{R}^S_+$ not in the independent set polytope of M, an inequality among (40.2) most violated by x. That is, to find $U \subseteq S$ maximizing $x(U) - r_M(U)$.

The following theorem implies a min-max relation for this (Edmonds [1970b]):

Theorem 40.3. *Let $M = (S, \mathcal{I})$ be a matroid and let $x \in \mathbb{R}^S_+$. Then*

(40.18) $$\max\{z(S) \mid z \in P_{\text{independent set}}(M), z \leq x\} = \min\{r_M(U) + x(S \setminus U) \mid U \subseteq S\}.$$

Proof. The inequality $\leq$ in (40.18) follows from

(40.19) $$z(S) = z(U) + z(S \setminus U) \leq r_M(U) + x(S \setminus U).$$

To see equality, let z attain the maximum. Then for each $s \in S$ with $z_s < x_s$ there exists a $U \subseteq S$ with $s \in U$ and $z(U) = r_M(U)$ (otherwise we can increase z_s). Now the collection of sets $U \subseteq S$ satisfying $z(U) = r_M(U)$ is closed under taking unions (and intersections), since if $z(T) = r_M(T)$ and $z(U) = r_M(U)$, then

(40.20) $$z(T \cup U) = z(T) + z(U) - z(T \cap U) \geq r_M(T) + r_M(U) - r_M(T \cap U) \geq r_M(T \cup U).$$

Hence there exists a $U \subseteq S$ such that $z(U) = r_M(U)$ and such that U contains each $s \in S$ with $z_s < x_s$. Hence:

(40.21) $$z(S) = z(U) + z(S \setminus U) = r_M(U) + x(S \setminus U),$$

giving (40.18). ∎

Cunningham [1984] showed that from an independence testing oracle for a matroid one can derive a strongly polynomial time algorithm to find for any given vector x, a maximum violated inequality for the independent set polytope.

More strongly, Cunningham showed that one can solve the following problem in strongly polynomial time:

(40.22) given: a matroid $M = (S, \mathcal{I})$, by an independence testing oracle, and an $x \in \mathbb{Q}_+^S$;
find: a $z \in P_{\text{independent set}}(M)$ with $z \leq x$ maximizing $z(S)$, with a decomposition of z as convex combination of incidence vectors of independent sets, and a subset U of S satisfying $z(S) = r_M(U) + x(S \setminus U)$.

By (40.18), the set U certifies that z maximizes $z(S)$. In the algorithm for (40.22), Cunningham utilized the 'consistent breadth-first search' based on lexicographic order, given by Schönsleben [1980] and Lawler and Martel [1982a].

To prove Cunningham's result, we first show two lemmas. The first lemma is used only to prove the second lemma. As in Section 39.9, we define for any independent set I of a matroid $M = (S, \mathcal{I})$:

(40.23) $$A(I) := \{(y, z) \mid y \in I, z \in S \setminus I, I - y + z \in \mathcal{I}\}.$$

Lemma 40.4α. *Let $M = (S, \mathcal{I})$ be a matroid and let $I \in \mathcal{I}$. Let $(s, t) \in A(I)$, define $I' := I - s + t$, and let $(u, v) \in A(I') \setminus A(I)$. Then $t = u$ or $(u, t) \in A(I)$, and $s = v$ or $(s, v) \in A(I)$.*

Proof. By symmetry, it suffices to show that $t = u$ or $(u, t) \in A(I)$ (as we may assume that I is a base, and hence the second part follows by duality). We can assume that $t \neq u$. Then $t \neq v$, since $v \notin I' = I - s + t$, as $(u, v) \in A(I')$.

If $v = s$, then $I - u + t = I - u - s + t + v = I' - u + v \in \mathcal{I}$ and hence $(u, t) \in A(I)$. If $v \neq s$, then $I - u \in \mathcal{I}$ and $I - u - s + t + v \in \mathcal{I}$, and therefore $I - u + t \in \mathcal{I}$ or $I - u + v \in \mathcal{I}$; that is, $(u, v) \in A(I)$ or $(u, t) \in A(I)$. ∎

Lemma 40.4β. *Let $M = (S, \mathcal{I})$ be a matroid and let q be a new element. For any $I \in \mathcal{I}$, define*

(40.24) $$\widetilde{A}(I) := \{(u, v) \mid u \in I + q, v \in S \setminus I, I - u + v \in \mathcal{I}\}.$$

Let $(s, t) \in A(I)$, define $I' := I - s + t$, and let $(u, v) \in \widetilde{A}(I') \setminus \widetilde{A}(I)$. Then $t = u$ or $(u, t) \in \widetilde{A}(I)$, and $s = v$ or $(s, v) \in \widetilde{A}(I)$.

Proof. Let $\widetilde{\mathcal{I}} := \{J \subseteq S + q \mid J - q \in \mathcal{I}\}$. Then the present lemma follows from Lemma 40.4α applied to the matroid $(S + q, \widetilde{\mathcal{I}})$. ∎

Now we can derive Cunningham's result:

Theorem 40.4. *Problem* (40.22) *is solvable in strongly polynomial time.*

Proof. We keep a vector $z \leq x$ in the independent set polytope of M and a decomposition

(40.25) $$z = \sum_{i=1}^{k} \lambda_i \chi^{I_i},$$

with $I_1, \ldots, I_k \in \mathcal{I}$, $\lambda_1, \ldots, \lambda_k > 0$, and $\sum_i \lambda_i = 1$. Initially $z := \mathbf{0}$, $k := 1$, $I_1 := \emptyset$, $\lambda_1 := 1$.

Let

(40.26) $$T := \{s \in S \mid z_s < x_s\}.$$

Let q be a new element. For each i, define $\widetilde{A}(I_i)$ as in (40.24), and let $D = (S + q, A)$ be the directed graph with

(40.27) $$A := \widetilde{A}(I_1) \cup \cdots \cup \widetilde{A}(I_k).$$

Fix an arbitrary linear order of the elements of $S + q$, by setting $S + q = \{1, \ldots, n\}$.

Case 1: D has no $q - T$ path. Let U be the set of $s \in S$ for which D has an $s - T$ path. As $T \subseteq U$, we know $z(S \setminus U) = x(S \setminus U)$. Also, as no arc of D enters U, we have $|U \cap I_i| = r_M(U)$ for all i, implying

(40.28) $$z(U) = \sum_{i=1}^{k} \lambda_i |U \cap I_i| = \sum_{i=1}^{k} \lambda_i r_M(U) = r_M(U).$$

Hence $z(S) = r_M(U) + x(S \setminus U)$ as required.

Case 2: D has a $q-T$ path. For each $v \in S+q$, let $d(v)$ denote the distance in D from q to v (set to ∞ if no $q-v$ path exists). Choose a $t \in T$ with $d(t)$ finite and maximal, and among these t we choose the largest t. Let $(s,t) \in A$, with $d(s) = d(t) - 1$, and s largest. We can assume that $(s,t) \in \widetilde{A}(I_1)$. Let

(40.29) $\alpha := \min\{x_t - z_t, \lambda_1\}$

and define z' by

(40.30) $z' := z + \alpha(\chi^t - \chi^s)$ if $s \neq q$, and $z' := z + \alpha\chi^t$ if $s = q$.

Let $I'_1 := I_1 - s + t$ (so $I'_1 = I_1 + t$ if $s = q$).

Then

(40.31) $$z' = \alpha\chi^{I'_1} + (\lambda_1 - \alpha)\chi^{I_1} + \sum_{i=2}^{k} \lambda_i \chi^{I_i}.$$

If $\alpha = \lambda_1$, we delete the second term. We obtain a decomposition of z' as a convex combination of at most $k+1$ independent sets, and we can iterate.

Running time. We show that the number of iterations is at most $|S|^9$. Consider any iteration. Let d' and A' be the objects d and A of the next iteration. We first show:

(40.32) for each $v \in S+q$: $d'(v) \geq d(v)$.

To show this, we can assume that $d'(v) < \infty$. We show (40.32) by induction on $d'(v)$, the case $d'(v) = 0$ being trivial (as it means $v = q$). Assume $d'(v) > 0$. Let u be such that $(u,v) \in A'$ and $d'(u) = d'(v) - 1$. By induction we know $d'(u) \geq d(u)$.

If $(u,v) \in A$, then $d(v) \leq d(u) + 1 \leq d'(u) + 1 = d'(v)$, as required. If $(u,v) \notin A$, then $(u,v) \in \widetilde{A}(I'_1)$ and $(u,v) \notin \widetilde{A}(I_1)$. By Lemma 40.4$\beta$, $t = u$ or $(u,t) \in \widetilde{A}(I_1)$, and $s = v$ or $(s,v) \in \widetilde{A}(I_1)$. Hence

(40.33) $d(v) \leq d(s) + 1 = d(t) \leq d(u) + 1 \leq d'(u) + 1 = d'(v)$.

So $d(v) \leq d'(v)$. This shows (40.32).

Let β be the number of $j = 1, \ldots, k$ with $(s,t) \in \widetilde{A}(I_j)$. Let T', t', s', and β' be the objects T, t, s, β in the next iteration. We show:

(40.34) if $d'(v) = d(v)$ for each $v \in S+q$, then $(d'(t'), t', s', \beta')$ is lexicographically less than $(d(t), t, s, \beta)$.

Indeed, if $\alpha = x_t - z_t$, then $T' = T - t + s$ or $T' = T - t$. So $d'(t') < d(t)$, or $d'(t') = d(t)$ and $t' < t$. If $\alpha < x_t - z_t$, then $T' = T + s$ or $T' = T$. Moreover, $\alpha = \lambda_1$, so I_1 has been omitted from the convex combination. So, as $t \in T'$ and $d(s) < d(t)$, we know that $t' = t$ and $d'(t') = d(t)$. As $t \in I'_1$, we know $(s',t) \notin \widetilde{A}(I'_1)$. Hence, as $(s',t) \in A'$, we have $(s',t) \in \widetilde{A}(I_j)$ for some $j = 2, \ldots, k$. Hence $(s',t) \in A$. By the choice of s, we know $s' \leq s$. If $s' < s$, we have (40.34), so assume $s' = s$. Then $\beta' = \beta - 1$, as $(s,t) \notin \widetilde{A}(I'_1)$. This proves (40.34).

The number k of independent sets in the decomposition grows by 1 if $\alpha = x_t - z_t < \lambda_1$. In that case, $d'(v) = d(v)$ for each $v \in S + q$ (by (40.32), as $A' \supseteq A$). Moreover, $d'(t') < d(t)$ or $t' < t$ (since $T' \subseteq T - t + s$). So k does not exceed $|S|^4$, and hence β is at most $|S|^4$. Concluding, the number of iterations is at most $|S|^9$. ▮

With Gaussian elimination, we can reduce the number k in each iteration to at most $|S|$ (by Carathéodory's theorem). Incorporating this reduces the number of iterations to $|S|^6$.

Theorem 40.4 immediately implies that one can test if a given vector belongs to the independent set polytope of a matroid:

Corollary 40.4a. *Given a matroid $M = (S, \mathcal{I})$ by an independence testing oracle and an $x \in \mathbb{Q}^S$, one can test in strongly polynomial time if x belongs to $P_{\text{independent set}}(M)$, and if so, decompose x as a convex combination of incidence vectors of independent sets.*

Proof. Directly from Theorem 40.4. ▮

One can derive a similar result for the spanning set polytope:

Corollary 40.4b. *Given a matroid $M = (S, \mathcal{I})$ by an independence testing oracle and an $x \in \mathbb{Q}^S$, one can test in strongly polynomial time if x belongs to $P_{\text{spanning set}}(M)$, and if so, decompose x as a convex combination of incidence vectors of spanning sets.*

Proof. x belongs to the spanning set polytope of M if and only if $\mathbf{1} - x$ belongs to the independent set polytope of the dual matroid M^*. Also convex combinations of spanning sets of M and independent sets of M^* transfer to each other by this operation. Since $r_{M^*}(U) = |U| + r_M(S \setminus U) - r_M(S)$ for each $U \subseteq S$, also an independence testing oracle for M^* is easily obtained from one for M. ▮

The theorem also implies that the following *most violated inequality problem* can be solved in strongly polynomial time:

(40.35) given: a matroid $M = (S, \mathcal{I})$ by an independence testing oracle, and a vector $x \in \mathbb{Q}^S$;
find: a subset U of S minimizing $r_M(U) - x(U)$.

Corollary 40.4c. *The most violated inequality problem can be solved in strongly polynomial time.*

Proof. Any negative component of x can be reset to 0, as this does not change the problem. So we can assume that $x \geq \mathbf{0}$. Then by Theorem 40.4 we can find a $U \subseteq S$ minimizing $r_M(U) + x(S \setminus U)$ in strongly polynomial time. This U is as required. ∎

40.3a. Facets and adjacency on the independent set polytope

Let $M = (S, \mathcal{I})$ be a matroid, with rank function r. Trivially, the independent set polytope P of M is full-dimensional if and only if M has no loops. If P is full-dimensional there is a unique minimal collection of linear inequalities defining P (up to scalar multiplication), which corresponds to the facets of P. Edmonds [1970b] found that this collection is given by the following theorem. Recall that a subset F of S is called a *flat* if for all s in $S \setminus F$ one has $r(F + s) > r(F)$. A subset F is called *inseparable* if there is no partition of F into nonempty sets F_1 and F_2 with $r(F) = r(F_1) + r(F_2)$. Then:

Theorem 40.5. *If M is loopless, the following is a minimal system for the independent set polytope of M:*

(40.36) (i) $x_s \geq 0$ $(s \in S)$,
(ii) $x(F) \leq r(F)$ (F is a nonempty inseparable flat).

Proof. As M is loopless, the independent set polytope of M is full-dimensional. It is easy to see that (40.36) determines the independent set polytope, as any other inequality $x(U) \leq r(U)$ is implied by the inequalities $x(F_i) \leq r(F_i)$, where $F_1, \ldots, F_t$ is a maximal partition of $F := \text{span}_M(U)$ such that $r(F_1) + \cdots + r(F_t) = r(F)$.

The irredundancy of collection (40.36) can be seen as follows. Each inequality $x_s \geq 0$ is irredundant, since the vector $-\chi^s$ satisfies all other inequalities.

We show that also the inequalities (40.36)(ii) are irredundant, by showing that for any two nonempty nonseparable flats T, U there exists a base I of T with $|I \cap U| < r(U)$ (implying that the face determined by T is contained in no (other) facet).

To show this, let I be a base of T with $|I \cap (T \setminus U)| = r(T \setminus U)$. Suppose $|I \cap U| = r(U)$. Then

(40.37) $$r(U) \geq r(T \cap U) \geq r(T) - r(T \setminus U) = |I \cap U| = r(U).$$

Hence we have equality throughout. This implies (as T is inseparable) that $T \setminus U = \emptyset$ or $T \cap U = \emptyset$, and that $r(U) = r(T \cap U)$. If $T \setminus U = \emptyset$, then $T \subset U$, and hence (as T is a flat) $r(U) > r(T) \geq r(T \cap U)$, a contradiction. If $T \cap U = \emptyset$, then $r(U) = r(T \cap U) = 0$, implying that $U = \emptyset$ (as M has no loops), again a contradiction. ∎

It follows that the base polytope, which is the face $\{x \in P \mid x(S) = r(S)\}$ of P, has dimension $|S| - 1$ if and only if S is inseparable (that is, the matroid is *connected*).

As for adjacency of vertices of the independent set polytope, we have:

Theorem 40.6. *Let $M = (S, \mathcal{I})$ be a loopless matroid and let I and J be distinct independent sets. Then χ^I and χ^J are adjacent vertices of the independent set* *polytope of* M *if and only if* $|I \triangle J| = 1$*, or* $|I \setminus J| = |J \setminus I| = 1$ *and* $r_M(I \cup J) = |I| = |J|$.

Proof. To see sufficiency, note that the condition implies that I and J are the only two independent sets with incidence vector x satisfying $x(I \cap J) = r_M(I \cap J)$, $x_s = 0$ for $s \notin I \cup J$, and (if $|I \triangle J| = 2$) $x(I \cup J) = r_M(I \cup J)$. Hence I and J are adjacent.

To see necessity, assume that χ^I and χ^J are adjacent. If I is not a base of $I \cup J$, then $I + j$ is independent for some $j \in J \setminus I$. Hence

(40.38) $$\tfrac{1}{2}(\chi^I + \chi^J) = \tfrac{1}{2}(\chi^{I+j} + \chi^{J-j}),$$

implying (as χ^I and χ^J are adjacent) that $I+j = J$ and $J-j = I$, that is $|I \triangle J| = 1$.

So we can assume that I and J are bases of $I \cup J$. Choose $i \in I \setminus J$. By Theorem 39.12, there is a $j \in J \setminus I$ such that $I - i + j$ and $J - j + i$ are bases of $I \cup J$. Then

(40.39) $$\tfrac{1}{2}(\chi^I + \chi^J) = \tfrac{1}{2}(\chi^{I-i+j} + \chi^{J-j+i}),$$

implying (as χ^I and χ^J are adjacent) that $I - i + j = J$ and $J - j + i = I$, that is we have the second alternative in the condition. ∎

More on the combinatorial structure of the independent set polytope can be found in Naddef and Pulleyblank [1981a].

40.3b. Further notes

Prodon [1984] showed that the separation problem for the independent set polytope of a matching matroid can be solved by finding a minimum-capacity cut in an auxiliary directed graph.

Frederickson and Solis-Oba [1997,1998] gave strongly polynomial-time algorithm for measuring the sensitivity of the minimum weight of a base under perturbing the weight. (Related analysis was given by Libura [1991].)

Narayanan [1995] described a rounding technique for the independent set polytope membership problem, leading to an $O(n^3r^2)$-time algorithm, where n is the size of the underlying set of the matroid and r is the rank of the matroid.

A strongly polynomial-time algorithm maximizing certain convex objective functions over the bases was given by Hassin and Tamir [1989].

For studies of structures where the greedy algorithm applies if condition (39.1)(i) is deleted, see Faigle [1979,1984b], Hausmann, Korte, and Jenkyns [1980], Korte and Lovász [1983,1984a,1984b,1984c,1985a,1985b,1989], Bouchet [1987a], Goecke [1988], Dress and Wenzel [1990], Korte, Lovász, and Schrader [1991], Helman, Moret, and Shapiro [1993], and Faigle and Kern [1996].

Chapter 41

Matroid intersection

Edmonds discovered that matroids have even more algorithmic power than just that of the greedy method. He showed that there exist efficient algorithms also for *intersections* of matroids. That is, a maximum-weight common independent set in *two* matroids can be found in strongly polynomial time. Edmonds also found good min-max characterizations for matroid intersection.

Matroid intersection yields a motivation for studying matroids: we may apply it to two matroids from different classes of examples of matroids, and thus we obtain methods that exceed the bounds of any particular class.

We should note here that if $M_1 = (S, \mathcal{I}_1)$ and $M_2 = (S, \mathcal{I}_2)$ are matroids, then $(S, \mathcal{I}_1 \cap \mathcal{I}_2)$ need not be a matroid. (An example with $|S| = 3$ is easy to construct.)

Moreover, the problem of finding a maximum-size common independent set in *three* matroids is NP-complete (as finding a Hamiltonian circuit in a directed graph is a special case; also, finding a common transversal of three partitions is a special case).

41.1. Matroid intersection theorem

Let $M_1 = (S, \mathcal{I}_1)$ and $M_2 = (S, \mathcal{I}_2)$ be two matroids, on the same set S. Consider the collection $\mathcal{I}_1 \cap \mathcal{I}_2$ of *common independent sets*. The pair $(S, \mathcal{I}_1 \cap \mathcal{I}_2)$ is generally *not* a matroid again.

Edmonds [1970b] showed the following formula, for which he gave two proofs — one based on linear programming duality and total unimodularity (see the proof of Theorem 41.12 below), and one reducing it to the matroid union theorem (see Corollary 42.1a and the remark thereafter). We give the direct proof implicit in Brualdi [1971e].

Theorem 41.1 (matroid intersection theorem)**.** *Let $M_1 = (S, \mathcal{I}_1)$ and $M_2 = (S, \mathcal{I}_2)$ be matroids, with rank functions r_1 and r_2, respectively. Then the maximum size of a set in $\mathcal{I}_1 \cap \mathcal{I}_2$ is equal to*

(41.1) $$\min_{U \subseteq S} (r_1(U) + r_2(S \setminus U)).$$

Proof. Let k be equal to (41.1). It is easy to see that the maximum is not more than k, since for any common independent set I and any $U \subseteq S$:

$$(41.2) \qquad |I| = |I \cap U| + |I \setminus U| \leq r_1(U) + r_2(S \setminus U).$$

We prove equality by induction on $|S|$, the case $|S| \leq 1$ being trivial. So assume that $|S| \geq 2$.

If minimum (41.1) is attained only by $U = S$ or $U = \emptyset$, choose $s \in S$. Then $r_1(U) + r_2(S \setminus (U \cup \{s\})) \geq k$ for each $U \subseteq S \setminus \{s\}$, since otherwise both U and $U \cup \{s\}$ would attain (41.1), whence $\{U, U \cup \{s\}\} = \{\emptyset, S\}$, contradicting the fact that $|S| \geq 2$. Hence, by induction, $M_1 \setminus s$ and $M_2 \setminus s$ have a common independent set of size k, implying the theorem.

So we can assume that (41.1) is attained by some U with $\emptyset \neq U \neq S$. Then $M_1|U$ and $M_2 \cdot U$ have a common independent set I of size $r_1(U)$. Otherwise, by induction, there exists a subset T of U with

$$(41.3) \qquad r_1(U) > r_{M_1|U}(T) + r_{M_2 \cdot U}(U \setminus T) = r_1(T) + r_2(S \setminus T) - r_2(S \setminus U),$$

contradicting the fact that U attains (41.1). Similarly, $M_1 \cdot (S \setminus U)$ and $M_2|(S \setminus U)$ have a common independent set J of size $r_2(S \setminus U)$.

Now $I \cup J$ is a common independent set of M_1 and M_2. Indeed, $I \cup J$ is independent in M_1, as I is independent in $M_1|U$ and J is independent in $M_1 \cdot (S \setminus U) = M_1/U$ (cf. (39.10)). Similarly, $I \cup J$ is independent in M_2. As $|I \cup J| = r_1(U) + r_2(S \setminus U)$, this proves the theorem. ∎

This implies a characterization of the existence of a common base in two matroids:

Corollary 41.1a. *Let $M_1 = (S, \mathcal{I}_1)$ and $M_2 = (S, \mathcal{I}_2)$ be matroids, with rank functions r_1 and r_2, respectively, such that $r_1(S) = r_2(S)$. Then M_1 and M_2 have a common base if and only if $r_1(U) + r_2(S \setminus U) \geq r_1(S)$ for each $U \subseteq S$.*

Proof. Directly from Theorem 41.1. ∎

It is easy to derive from the matroid intersection theorem a similar min-max relation for the minimum size of a common spanning set:

Corollary 41.1b. *Let $M_1 = (S, \mathcal{I}_1)$ and $M_2 = (S, \mathcal{I}_2)$ be matroids, with rank functions r_1 and r_2, respectively. Then the minimum size of a common spanning set of M_1 and M_2 is equal to*

$$(41.4) \qquad \max_{U \subseteq S} (r_1(S) - r_1(U) + r_2(S) - r_2(S \setminus U)).$$

Proof. The minimum is equal to the minimum of $|B_1 \cup B_2|$ where B_1 and B_2 are bases of M_1 and M_2 respectively. Hence the minimum is equal to $r_1(S) + r_2(S)$ minus the maximum of $|B_1 \cap B_2|$ over such B_1, B_2. This last maximum is characterized in the matroid intersection theorem, yielding the present corollary. ∎

The following result of Rado [1942] (a generalization of Hall's marriage theorem (Theorem 22.1), and therefore sometimes called the Rado-Hall theorem) may be derived from the matroid intersection theorem, applied to M and the transversal matroid M_2 induced by $\mathcal{X}$.

Corollary 41.1c (Rado's theorem)**.** *Let $M = (S, \mathcal{I})$ be a matroid, with rank function r, and let $\mathcal{X} = (X_1, \ldots, X_n)$ be a family of subsets of S. Then $\mathcal{X}$ has a transversal which is independent in M if and only if*

$$(41.5) \qquad r(\bigcup_{i \in I} X_i) \geq |I|$$

for each $I \subseteq \{1, \ldots, n\}$.

Proof. Let r_2 be the rank function of the transversal matroid M_2 induced by $\mathcal{X}$. By the matroid intersection theorem, M and M_2 have a common independent set of size n if and only if

$$(41.6) \qquad r(U) + r_2(S \setminus U) \geq n \text{ for each } U \subseteq S.$$

Now for each $T \subseteq S$ one has (by Kőnig's matching theorem (cf. Corollary 22.2a)):

$$(41.7) \qquad r_2(T) = \min_{I \subseteq \{1,\ldots,n\}} (\big|\bigcup_{i \in I} X_i \cap T\big| + n - |I|).$$

So (41.6) is equivalent to:

$$(41.8) \qquad r(U) + \big|\bigcup_{i \in I} X_i \setminus U\big| + n - |I| \geq n$$

for all $U \subseteq S$ and $I \subseteq \{1, \ldots, n\}$. We can assume that $U = \bigcup_{i \in I} X_i$, since replacing U by $\bigcup_{i \in I} X_i$ does not increase the left-hand side in (41.8). So the condition is equivalent to (41.5), proving the corollary. ∎

Notes. Mirsky [1971a] gave an alternative proof of Rado's theorem. Welsh [1970] showed that, in turn, Rado's theorem implies the matroid intersection theorem. Las Vergnas [1970] gave an extension of Rado's theorem. Rado [1942] (and also Welsh [1971]) showed that Rado's theorem in fact characterizes matroids. Perfect [1969a] generalized Rado's theorem to characterizing the maximum size of an independent partial transversal. Related results are in Perfect [1971].

41.1a. Applications of the matroid intersection theorem

In this section we mention a number of applications of the matroid intersection theorem. Further applications will be given in the next chapter on matroid union.

Kőnig's theorems. Let $G = (V, E)$ be a bipartite graph, with colour classes U_1 and U_2. For $i = 1, 2$, let $M_i = (E, \mathcal{I}_i)$ be the matroid with $F \subseteq E$ independent if and only if each vertex in U_i is covered by at most one edge in F.

So M_1 and M_2 are partition matroids. The common independent sets in M_1 and M_2 are the matchings in G, and the common spanning sets are the edge covers in G. For $i = 1, 2$ and $F \subseteq E$, the rank $r_i(F)$ of F in M_i is equal to the number of vertices in U_i covered by F.

By the matroid intersection theorem, the maximum size of a matching in G is equal to the minimum of $r_1(F) + r_2(E \setminus F)$ taken over $F \subseteq E$. This last is equal to the minimum size of a vertex cover in G. So we have Kőnig's matching theorem (Theorem 16.2).

Similarly, by Corollary 41.1b, the minimum size of an edge cover in G (assuming G has no isolated vertices), is equal to the maximum of $|V| - r_1(F) - r_2(E \setminus F)$ taken over $F \subseteq E$. This last is equal to the maximum size of a stable set in G. So we have the Kőnig-Rado edge cover theorem (Theorem 19.4).

Common transversals. Let $\mathcal{X} = (X_1, \ldots, X_m)$ and $\mathcal{Y} = (Y_1, \ldots, Y_m)$ be families of subsets of a finite set S. Then the matroid intersection theorem implies Theorem 23.1 of Ford and Fulkerson [1958c]: $\mathcal{X}$ and $\mathcal{Y}$ have a common transversal if and only if

(41.9) $$|X_I \cap Y_J| \geq |I| + |J| - m$$

for all subsets I and J of $\{1, \ldots, m\}$, where $X_I := \bigcup_{i \in I} X_i$ and $Y_J := \bigcup_{j \in J} Y_j$.

To see this, let M_1 and M_2 be the transversal matroids induced by $\mathcal{X}$ and $\mathcal{Y}$ respectively, with rank functions r_1 and r_2 say. So $\mathcal{X}$ and $\mathcal{Y}$ have a common transversal if and only if M_1 and M_2 have a common independent set of size m. By Theorem 41.1, this last holds if and only if $r_1(Z) + r_2(S \setminus Z) \geq m$ for each $Z \subseteq S$. Using Kőnig's matching theorem, this is equivalent to:

(41.10) $$\min_{I \subseteq \{1,\ldots,m\}} (m - |I| + |X_I \cap Z|) + \min_{J \subseteq \{1,\ldots,m\}} (m - |J| + |Y_J \setminus Z|) \geq m$$

for each $Z \subseteq S$. Equivalently, for all $I, J \subseteq \{1, \ldots, m\}$:

(41.11) $$\min_{Z \subseteq S} (m - |I| + |X_I \cap Z| + m - |J| + |Y_J \setminus Z|) \geq m.$$

As this minimum is attained by $Z := Y_J$, this is equivalent to (41.9).

Coloured trees. Let $G = (V, E)$ be a graph and let the edges of G be coloured with k colours. That is, we have partitioned E into sets $E_1, \ldots, E_k$, called *colours*. Then there exists a spanning tree with all edges coloured differently if and only if $G - F$ has at most $t + 1$ components, for any union F of t colours, for any $t \geq 0$. This follows from the matroid intersection theorem applied to the cycle matroid $M(G)$ of G and the partition matroid N induced by $E_1, \ldots, E_k$.

Indeed, $M(G)$ and N have a common independent set of size $|V| - 1$ if and only if $r_{M(G)}(E \setminus F) + r_N(F) \geq |V| - 1$ for each $F \subseteq E$. Now $r_N(F)$ is equal to the number of E_i intersecting F. So we can assume that F is equal to the union of t of the E_i, with $t := r_N(F)$. Moreover, $r_{M(G)}(E \setminus F)$ is equal to $|V| - \kappa(G - F)$, where $\kappa(G - F)$ is the number of components of $G - F$. So the requirement is that $|V| - \kappa(G - F) + t \geq |V| - 1$. In other words, $\kappa(G - F) \leq t + 1$.

Detachments. The following is a special case of a theorem of Nash-Williams [1985], which he derived from the matroid intersection theorem — in fact it is a consequence of the result on coloured trees given above.

Let $G = (V, E)$ be a graph and let $b : V \longrightarrow \mathbb{Z}_+$. Call a graph $\widetilde{G} = (\widetilde{V}, \widetilde{E})$ a *b-detachment* of G if there is a function $\phi : \widetilde{V} \longrightarrow V$ such that $|\phi^{-1}(v)| = b(v)$ for each $v \in V$, and such that there is a one-to-one function $\psi : \widetilde{E} \longrightarrow E$ with $\psi(e) = \{\phi(u), \phi(v)\}$ for each edge $e = uv$ of $\widetilde{G}$.

Then there exists a connected b-detachment if and only if

(41.12) $\quad b(U) + \kappa(G - U) \leq |E_U| + 1$ for each $U \subseteq V$,

where $\kappa(G')$ denotes the number of components of graph G' and where E_U denotes the set of edges intersecting U.

To see this, let $H = (\widetilde{V}, E')$ be the graph obtained from G by replacing each vertex v by $b(v)$ new vertices, and by connecting for each edge $e = uv$ of G, the $b(u)$ new vertices associated with u with the $b(v)$ new vertices associated with v. We assign to these $b(u)b(v)$ edges the 'colour' e.

Then there exists a connected b-detachment if and only if H has a spanning tree in which all edges have a different colour. By the previous example, such a spanning tree exists if and only if for each $F \subseteq E$, deleting from H the edges with colour in F gives a graph H' with at most $|F| + 1$ components.

Now the number of components of H' is equal to the $\kappa(G - F) + b(I_F) - |I_F|$, where I_F denotes the set of isolated (hence loopless) vertices of $G - F$. So the condition is equivalent to: $\kappa(G - F) - |F| + b(I_F) - |I_F| \leq 1$. As $\kappa(G - F) - |F|$ does not decrease by removing edges from F, we can assume that F is equal to the set of edges incident with I_F. So F is determined by $U := I_F$, namely $F = E_U$. Then $\kappa(G - F) - |I_F| = \kappa(G - U)$. So the condition is equivalent to (41.12).

41.1b. Woodall's proof of the matroid intersection theorem

P.D. Seymour attributed the following proof of the matroid intersection theorem to D.R. Woodall (cf. Seymour [1976a]):

Let k be the value of (41.1). Let $x \in S$ be such that $r_1(\{x\}) = r_2(\{x\}) = 1$. (If no such x exists the theorem is trivial, as in that case the minimum is 0.) Let $Y := S \setminus \{x\}$. Now we may assume that the restrictions $M_1 \setminus x$ and $M_2 \setminus x$ have no common independent set of size k. So, by induction,

(41.13) $\quad r_1(A_1) + r_2(A_2) \leq k - 1,$

for some partition A_1, A_2 of Y. Moreover, the contractions M_1/x and M_2/x have no common independent set of size $k-1$ (otherwise we can add x to obtain a common independent set of size k for M_1 and M_2). So, by induction,

(41.14) $\quad r_1(B_1 \cup \{x\}) - 1 + r_2(B_2 \cup \{x\}) - 1 \leq k - 2$

(cf. (39.9) above), for some partition B_1, B_2 of Y. However,

(41.15)
$$\begin{aligned} r_1(A_1 \cap B_1) + r_1(A_1 \cup B_1 \cup \{x\}) &\leq r_1(A_1) + r_1(B_1 \cup \{x\}), \\ r_2(A_2 \cap B_2) + r_2(A_2 \cup B_2 \cup \{x\}) &\leq r_2(A_2) + r_2(B_2 \cup \{x\}), \end{aligned}$$

by the submodularity (cf. (39.38)(ii)) of the rank functions. Moreover, by the definition of k,

(41.16)
$$\begin{aligned} k &\leq r_1(A_1 \cap B_1) + r_2(A_2 \cup B_2 \cup \{x\}), \\ k &\leq r_1(A_1 \cup B_1 \cup \{x\}) + r_2(A_2 \cap B_2), \end{aligned}$$

as $A_1 \cap B_1, A_2 \cup B_2 \cup \{x\}$ and $A_1 \cup B_1 \cup \{x\}, A_2 \cap B_2$ form partitions of S. Adding the inequalities in (41.13), (41.14), (41.15), and (41.16) gives a contradiction.

41.2. Cardinality matroid intersection algorithm

A maximum-size common independent set can be found in polynomial time. This result follows from the matroid union algorithm of Edmonds [1968], since (as Edmonds [1970b] and Lawler [1970] observed) cardinality matroid intersection can be reduced to matroid union.

We describe below the direct algorithm given by Aigner and Dowling [1971] and Lawler [1975], based on finding paths in auxiliary graphs. A different algorithm was given by Edmonds [1979].

Note that the examples given in Section 41.1a provide applications for the matroid intersection algorithm. We should note that in the algorithm we require that in any matroid $M = (S, \mathcal{I})$, we can test in polynomial time if any subset of S belongs to $\mathcal{I}$ — no explicit list of all sets in $\mathcal{I}$ is required. Thus complexity results are all relative to the complexity of testing independence. As such a membership testing algorithm exists in each example mentioned, we obtain polynomial-time algorithms for these special cases.

For any two matroids $M_1 = (S, \mathcal{I}_1)$ and $M_2 = (S, \mathcal{I}_2)$ and any $I \in \mathcal{I}_1 \cap \mathcal{I}_2$, we define a directed graph $D_{M_1,M_2}(I)$, with vertex set S, as follows. For any $y \in I, x \in S \setminus I$,

(41.17) (y, x) is an arc of $D_{M_1,M_2}(I)$ if and only if $I - y + x \in \mathcal{I}_1$,
(x, y) is an arc of $D_{M_1,M_2}(I)$ if and only if $I - y + x \in \mathcal{I}_2$.

These are all arcs of $D_{M_1,M_2}(I)$. So this graph is the union of the graphs $D_{M_1}(I)$ and the *reverse* of $D_{M_2}(I)$ defined in Section 39.9.

The following is the base for finding a maximum-size common independent set in two matroids.

Cardinality common independent set augmenting algorithm

input: matroids $M_1 = (S, \mathcal{I}_1)$ and $M_2 = (S, \mathcal{I}_2)$ and a set $I \in \mathcal{I}_1 \cap \mathcal{I}_2$;
output: a set $I' \in \mathcal{I}_1 \cap \mathcal{I}_2$ with $|I'| > |I|$ (if any).
description of the algorithm: Consider the sets

(41.18) $X_1 := \{x \in S \setminus I \mid I \cup \{x\} \in \mathcal{I}_1\}$,
$X_2 := \{x \in S \setminus I \mid I \cup \{x\} \in \mathcal{I}_2\}$.

Moreover, consider the directed graph $D_{M_1,M_2}(I)$ defined above. There are two cases.

Case 1: $D_{M_1,M_2}(I)$ has an $X_1 - X_2$ path P. (Possibly of length 0 if $X_1 \cap X_2 \neq \emptyset$.) We take a shortest such path P (that is, with a minimum number of arcs). Now output $I' := I \triangle VP$.

Case 2: $D_{M_1,M_2}(I)$ has no $X_1 - X_2$ path. Then I is a maximum-size common independent set. ▮

This finishes the description of the algorithm. The correctness of the algorithm is given by the following two theorems.

Theorem 41.2. *If Case 1 applies, then* $I' \in \mathcal{I}_1 \cap \mathcal{I}_2$.

Proof. Assume that Case 1 applies. By symmetry it suffices to show that I' belongs to $\mathcal{I}_1$.

Let P start at $z_0 \in X_1$. The arcs in P leaving I form the only matching in $D_{M_1}(I)$ with union equal to $VP - z_0$, since otherwise P would have a shortcut. Moreover, for each $z \in VP \setminus I$ with $z \neq z_0$, one has $I + z \notin \mathcal{I}_1$, since otherwise $z \in X_1$, and hence P would have a shortcut. So by Corollary 39.13a, I' belongs to $\mathcal{I}_1$. ∎

Theorem 41.3. *If Case 2 applies, then I is a maximum-size common independent set.*

Proof. As Case 2 applies, there is no $X_1 - X_2$ path in $D_{M_1,M_2}(I)$. Hence there is a subset U of S with $X_1 \cap U = \emptyset$ and $X_2 \subseteq U$, and such that no arc enters U. We show

(41.19) $$r_{M_1}(U) + r_{M_2}(S \setminus U) \leq |I|.$$

To this end, we first show

(41.20) $$r_{M_1}(U) \leq |I \cap U|.$$

Suppose that $r_{M_1}(U) > |I \cap U|$. Then there exists an x in $U \setminus I$ such that $(I \cap U) \cup \{x\} \in \mathcal{I}_1$. Since $I \cup \{x\} \notin \mathcal{I}_1$ (as $x \notin X_1$), there is a $y \in I \setminus U$ with $I - y + x \in \mathcal{I}_1$. But then $D_{M_1}(I)$ has an arc from y to x, contradicting the facts that $x \in U$ and $y \notin U$ and that no arc enters U.

This shows (41.20). Similarly, $r_{M_2}(S \setminus U) \leq |I \setminus U|$. Hence we have (41.19). So by the matroid intersection theorem, I is a maximum-size common independent set. ∎

Clearly, the running time of the algorithm is polynomially bounded, since we can construct the auxiliary directed graph $D_{M_1,M_2}(I)$ and find the path P (if any), in polynomial time. Therefore:

Theorem 41.4. *A maximum-size common independent set in two matroids can be found in polynomial time.*

Proof. Directly from the above, as we can find a maximum-size common independent set after applying at most $|S|$ times the common independent set augmenting algorithm. ∎

The algorithm also yields a proof of the matroid intersection theorem (Theorem 41.1 above): if the algorithm stops with set I, we obtain a set U for which (41.19) holds.

Notes. The above algorithm can be shown to take $O(n^2m(n+Q))$ time, where n is the maximum size of a common independent set, m is the size of the underlying set, and Q is the time needed to test if a given set is independent (in either matroid). Cunningham [1986] showed that if one chooses a shortest path as augmenting path, the sum of the lengths of all augmenting paths chosen is $O(n \log n)$, which gives an $O(n^{3/2}mQ)$-time algorithm. This algorithm extends several of the ideas behind the $O(n^{1/2}m)$ algorithm of Hopcroft, Karp, and Karzanov for cardinality bipartite matching (see Section 16.4). For more efficient algorithms, see Gabow and Tarjan [1984], Gusfield [1984], Gabow and Stallmann [1985], Frederickson and Srinivas [1989], Gabow and Xu [1989,1996], and Fujishige and Zhang [1995].

The problem of finding a maximum-size common independent set in *three* matroids is NP-complete, as finding a Hamiltonian circuit in a directed graph is a special case (as was observed by Held and Karp [1970]). Another special case is finding a common transversal of three collections of sets, which is also NP-complete (Theorem 23.16). In particular, the k-intersection problem can be reduced to the 3-intersection problem (cf. Lawler [1976b]).

Barvinok [1995] gave an algorithm for finding a maximum-size common independent set in k linear matroids, represented by given vectors over the rationals. The running time is linear in the cardinality of the underlying set and singly polynomial in the maximum rank of the matroids.

41.3. Weighted matroid intersection algorithm

Also a maximum-*weight* common independent set can be found in strongly polynomial time. This result was announced by Edmonds [1970b], who published an algorithm in Edmonds [1979]. An alternative algorithm (which we describe below) was announced by Lawler [1970] and described in Lawler [1975,1976b] — the correctness of this algorithm was proved by Krogdahl [1974,1976], using the results described in Section 39.9. A similar algorithm was described by Iri and Tomizawa [1976].

This algorithm is an extension of the cardinality matroid intersection algorithm given in Section 41.2. In each iteration, instead of finding a path P with a minimum number of arcs in $D_{M_1,M_2}(I)$, we will now require P to have minimum length with respect to some length function defined on $D_{M_1,M_2}(I)$.

To describe the algorithm, if matroids $M_1 = (S, \mathcal{I}_1)$ and $M_2 = (S, \mathcal{I}_2)$ and a weight function $w : S \to \mathbb{R}$ are given, call a set $I \in \mathcal{I}_1 \cap \mathcal{I}_2$ *extreme* if $w(J) \leq w(I)$ for each $J \in \mathcal{I}_1 \cap \mathcal{I}_2$ satisfying $|J| = |I|$.

Weighted common independent set augmenting algorithm

input: matroids $M_1 = (S, \mathcal{I}_1)$ and $M_2 = (S, \mathcal{I}_2)$, a weight function $w : S \to \mathbb{Q}$, and an extreme common independent set I;
output: an extreme common independent set I' with $|I'| = |I| + 1$ (if any).
description of the algorithm: Consider again the sets X_1 and X_2 and the directed graph $D_{M_1,M_2}(I)$ on S, as in the cardinality case.

For any $x \in S$ define the 'length' $l(x)$ of x by:

$$(41.21) \qquad l(x) := \begin{cases} w(x) & \text{if } x \in I, \\ -w(x) & \text{if } x \notin I. \end{cases}$$

The *length* of a path P, denoted by $l(P)$, is equal to the sum of the lengths of the vertices traversed by P.

Case 1: $D_{M_1,M_2}(I)$ has an $X_1 - X_2$ path P. We choose P such that $l(P)$ is minimal and such that (secondly) P has a minimum number of arcs among all minimum-length $X_1 - X_2$ paths. Set $I' := I \triangle VP$.

Case 2: $D_{M_1,M_2}(I)$ has no $X_1 - X_2$ path. Then there is no common independent set larger than I. ∎

This finishes the description of the algorithm. The correctness of the algorithm if Case 2 applies follows directly from Theorem 41.3. In order to show the correctness if Case 1 applies, we first prove the following basic property of the length function l.

Lemma 41.5α. *Let C be a directed circuit in $D_{M_1,M_2}(I)$ and let $t \in VC$. Define $J := I \triangle VC$. If $J \notin \mathcal{I}_1 \cap \mathcal{I}_2$, then there exists a directed circuit C' with $VC' \subset VC$ such that $l(VC') < 0$, or $l(VC') \leq l(VC)$ and $t \in VC'$.*

Proof. By symmetry we can assume that $J \notin \mathcal{I}_1$. Let N_1 and N_2 be the sets of arcs in C belonging to $D_{M_1}(I)$ and $D_{M_2}(I)$ respectively. As $J \notin \mathcal{I}_1$, there exists, by Theorem 39.13, a matching N_1' in $D_{M_1}(I)$ with union VC and with $N_1' \neq N_1$. Consider the directed graph $D = (VC, A)$ formed by the arcs in N_1, N_1' (taking arcs in $N_1 \cap N_1'$ parallel), and by the arcs in N_2 taking each of them twice (parallel). Then each vertex in VC is entered and left by exactly two arcs of D. Moreover, since $N_1' \neq N_1$, D contains a directed circuit C_1 with $VC_1 \subset VC$ (as N_1' contains a chord of C). As D is Eulerian, we can extend this to a decomposition of A into directed circuits $C_1, \ldots, C_k$. Then

$$(41.22) \qquad \chi^{VC_1} + \cdots + \chi^{VC_k} = 2 \cdot \chi^{VC}.$$

Since $VC_1 \neq VC$ we know that $VC_j = VC$ for at most one j. If, say $VC_k = VC$, then (41.22) implies that either $l(VC_j) < 0$ for some $j < k$ or $l(VC_j) \leq l(VC)$ for all $j < k$, implying the proposition.

Suppose next that $VC_j \neq VC$ for all j. If $l(VC_j) < 0$ for some $j \leq k$ we are done. So assume $l(VC_j) \geq 0$ for each $j \leq k$. We can assume that C_1 and C_2 traverse t. Then

$$(41.23) \qquad l(VC_1) + l(VC_2) \leq l(VC_1) + \cdots + l(VC_k) = 2l(VC).$$

Hence $l(VC_1) \leq l(VC)$ or $l(VC_2) \leq l(VC)$, and again we are done. ∎

This implies (Krogdahl [1976], Fujishige [1977a]):

Theorem 41.5. *Let $I \in \mathcal{I}_1 \cap \mathcal{I}_2$. Then I is extreme if and only if $D_{M_1,M_2}(I)$ has no directed circuit of negative length.*

Proof. To see necessity, suppose that $D_{M_1,M_2}(I)$ has a directed circuit C of negative length. Choose C with $|VC|$ minimal. Consider $J := I \triangle VC$. Since $w(J) = w(I) - l(C) > w(I)$, while $|J| = |I|$, we know that $J \notin \mathcal{I}_1 \cap \mathcal{I}_2$. Hence by Lemma 41.5$\alpha$, $D_{M_1,M_2}(I)$ has a negative-length directed circuit covering fewer than $|VC|$ vertices, contradicting our assumption.

To see sufficiency, consider a $J \in \mathcal{I}_1 \cap \mathcal{I}_2$ with $|J| = |I|$. By Corollary 39.12a, both $D_{M_1}(I)$ and $D_{M_2}(I)$ have a perfect matching on $I \triangle J$. These two matchings together form a vertex-disjoint union of a number of directed circuits $C_1, \ldots, C_t$. Then

(41.24) $$w(I) - w(J) = \sum_{j=1}^{t} l(VC_j) \geq 0,$$

implying $w(J) \leq w(I)$. So I is extreme. ∎

This theorem implies that we can find a shortest path P, in Case 1 of the algorithm, in strongly polynomial time (with the Bellman-Ford method). It also gives:

Theorem 41.6. *If Case 1 applies, I' is an extreme common independent set.*

Proof. We first show that $I' \in \mathcal{I}_1 \cap \mathcal{I}_2$. To this end, let t be a new element, and extend (for each $i = 1, 2$), M_i to a matroid $M_i' = (S + t, \mathcal{I}_i')$, where for each $T \subseteq S + t$:

(41.25) $T \in \mathcal{I}_i'$ if and only if $T - t \in \mathcal{I}_i$.

Note that $D_{M_1',M_2'}(I + t)$ arises from $D_{M_1,M_2}(I)$ by extending it with a new vertex t and adding arcs from t to each vertex in X_1, and from each vertex in X_2 to t.

Let P be the path found in the algorithm. Define

(41.26) $$w(t) := l(t) := -l(P).$$

As P is a shortest $X_1 - X_2$ path, this makes that $D_{M_1',M_2'}(I + t)$ has no negative-length directed circuit. Hence, by Theorem 41.5, $I + t$ is an extreme common independent set of M_1' and M_2'.

Let P run from $z_1 \in X_1$ to $z_2 \in X_2$. Extend P by the arcs (t, z_1) and (z_2, t) to a directed circuit C. So $J = (I + t) \triangle VC$. As P has a minimum number of arcs among all shortest $X_1 - X_2$ paths, and as $D_{M_1',M_2'}(I + t)$ has no negative-length directed circuits, by Lemma 41.5α we know that $J \in \mathcal{I}_1 \cap \mathcal{I}_2$.

Moreover, J is extreme, since $I + t$ is extreme and $w(J) = w(I + t)$. ∎

So the weighted common independent set augmenting algorithm is correct. It obviously has strongly polynomially bounded running time. Therefore:

Theorem 41.7. *A maximum-weight common independent set in two matroids can be found in strongly polynomial time.*

Proof. Starting with the extreme common independent set $I_0 := \emptyset$ we can find iteratively extreme common independent sets $I_0, I_1, \ldots, I_k$, where $|I_i| = i$ for $i = 0, \ldots, k$ and where I_k is a maximum-size common independent set. Taking one among $I_0, \ldots, I_k$ of maximum weight, we have a maximum-weight common independent set. ∎

The above algorithm gives a maximum-weight common independent set of size k, for each k. In particular, a maximum-weight common base can be found with the algorithm. Similarly for minimum-weight:

Theorem 41.8. *A minimum-weight common base in two matroids can be found in strongly polynomial time.*

Proof. The last extreme common independent set in the above algorithm is a maximum-weight common base. By flipping the signs of the weights, this can be turned into a minimum-weight common base algorithm. ∎

Notes. Frank [1981a] gave an $O(\tau n^3)$-time implementation of this algorithm, where τ is the time needed to test for any $I \in \mathcal{I}_i$ and any $s \in S$ whether or not $I \cup \{s\} \in \mathcal{I}_i$, and if not, to find a circuit of M_i contained in $I \cup \{s\}$.

Clearly, a maximum-weight common independent set need not be a common base, even if common bases exist and all weights are positive: Let $S = \{1, 2, 3\}$ and let M_i be the matroid on S with unique circuit $S \setminus \{i\}$ (for $i = 1, 2$). Define $w(1) := w(2) := 1$ and $w(3) := 3$. Then $\{3\}$ is the unique maximum-weight common independent set, while $\{1, 2\}$ is the unique common base.

41.3a. Speeding up the weighted matroid intersection algorithm

The algorithm described in Section 41.3 is strongly polynomial-time, since we can find a shortest path P in strongly polynomial time, as in each iteration the graph $D_{M_1,M_2}(I)$ has no negative-length directed circuit. Hence we can apply the Bellman-Ford method. To bound the running time, suppose that we can construct, for any $I \in \mathcal{I}_1 \cap \mathcal{I}_2$ the graph $D_{M_1,M_2}(I)$ in time T. Then any iteration can be done in time $O(T + n^3)$, where $n := |S|$.

We can improve this to $O(T + n \log n)$ as follows (Frank [1981a], Brezovec, Cornuéjols, and Glover [1986]). The idea is that, in each iteration, with the extreme common independent set I, we give a 'certificate' of extremity, by specifying a potential for the length function; that is, a function $p \in \mathbb{Q}^S$ satisfying

(41.27) $\qquad l(v) \geq p(v) - p(u)$

for each arc (u, v) of $D_{M_1,M_2}(I)$. By Theorem 41.5, such a potential certifies extremity of I. We call such a p a potential *for* I.

Having the potential, we can apply Dijkstra's method instead of the Bellman-Ford method, as with the potential we can transform the length function (if defined on arcs) to a nonnegative length function.

It is convenient to associate the following functions $w_1, w_2 : S \to \mathbb{R}$ to $p, w : S \to \mathbb{R}$:

(41.28) $w_1(v) = p(v)$ and $w_2(v) = w(v) - p(v)$ if $v \in I$,
$w_1(v) = w(v) + p(v)$ and $w_2(v) = -p(v)$ if $v \in S \setminus I$.

So $w = w_1 + w_2$. Then:

Theorem 41.9. *Let $I \in \mathcal{I}_1 \cap \mathcal{I}_2$ and let $p, w, w_1, w_2 : S \to \mathbb{R}$ satisfy* (41.28). *Then p is a potential for $D_{M_1,M_2}(I)$ if and only if for $i = 1, 2$ one has*

(41.29) *I maximizes $w_i(X)$ over all $J \in \mathcal{I}_i$ satisfying $|J| = |I|$.*

Proof. The theorem follows easily with Corollary 39.12b. Indeed, there is an arc (u, v) leaving I if and only if $I - u + v \in \mathcal{I}_1$. Then

(41.30) $w_1(v) \leq w_1(u) \iff l(v) \geq p(v) - p(u)$,

since $l(v) = -w(v) = -w_2(v) - w_1(v)$ and $-w_2(v) - w_1(u) = p(v) - p(u)$.

Similarly, there is an arc (u, v) entering I if and only if $I - v + u \in \mathcal{I}_2$. Then

(41.31) $w_2(v) \geq w_2(u) \iff l(v) \geq p(v) - p(u)$,

since $l(v) = w(v) = w_2(v) + w_1(v)$ and $w_2(u) + w_1(v) = p(v) - p(u)$. ∎

We trivially have a potential for $I := \emptyset$. Consider next an arbitrary iteration, with as input a common independent set I and a potential p for I. Construct $D_{M_1,M_2}(I)$ and l as before. Let P be an $X_1 - X_2$ path with $l(P)$ minimum, and, under this condition, with $|VP|$ minimum. (Using the potential described above, we can find P with Dijkstra's algorithm.) Let $I' := I \triangle VP$.

We now reset the potential p such that for any $v \in S$ with v reachable from X_1, $p(v)$ is equal to the distance from X_1 to v (= the minimum of $l(VQ)$ over all $X_1 - v$ paths Q in $D_{M_1,M_2}(I)$).

Let w_1 and w_2 satisfy (41.28) with respect to I, (the new) p, and w. Then:

Theorem 41.10. *w_1, w_2 satisfy* (41.29) *with respect to I'.*

Proof. Extend M_1 and M_2 to matroids $M_1' = (S + t, \mathcal{I}_1')$ and $M_2' = (S + t, \mathcal{I}_2')$ as in (41.25). Let P run from $z_1 \in X_1$ to $z_2 \in X_2$. Define $w(t) := l(t) := -l(P)$, $p(t) := 0$, $w_1(t) := 0$, and $w_2(t) := w(t)$. Now it suffices to show:

(41.32) (i) $w_i(I + t) = w_i(I')$ for $i = 1, 2$;
(ii) w_1, w_2 satisfy (41.29) with respect to M_1', M_2', and $I + t$.

Let C be the directed circuit obtained by extending P by the arcs (t, z_1) and (z_2, t). Now, since $I' = (I + t) \triangle VC$, to show (41.32), it suffices to show, for each arc (u, v):

(41.33) if (u, v) leaves $I + t$, then $w_1(v) \leq w_1(u)$, with equality if (u, v) is on C;
if (u, v) enters $I + t$, then $w_2(u) \leq w_2(v)$, with equality if (u, v) is on C.

Note that for each arc (u, v) of $D_{M_1',M_2'}(I + t)$ one has $p(v) \leq p(u) + l(v)$, with equality if (u, v) is on C. Hence, if (u, v) leaves $I + t$, then:

(41.34) $$w_1(v) = p(v) + w(v) = p(v) - l(v) \leq p(u) = w_1(u),$$

with equality if (u, v) is on C.

Similarly, if (u, v) enters $I + t$, then:

(41.35) $$w_2(v) = w(v) - p(v) = l(v) - p(v) \geq -p(u) = w_2(u),$$

with equality if (u, v) is on C. This proves (41.33). ∎

Using (41.28) and Theorem 41.9, we can obtain from w_1, w_2 a potential for I'. This implies:

Corollary 41.10a. *A maximum-weight common independent set can be found in time $O(k(T+n\log n))$, where $n := |S|$, k is the maximum size of a common independent set, and T is the time needed to find $D_{M_1,M_2}(I)$ for any common independent set I.*

Proof. Each iteration can be done in time $O(T + n\log n)$, since constructing the graph $D_{M_1,M_2}(I)$ takes T time, implying that there are $O(T)$ arcs. Hence, by Corollary 7.7a, a shortest $X_1 - X_2$ path P can be found in $O(T + n\log n)$ time. Hence I', and a potential for I' can be found in time $O(T + n\log n)$.

Since there are k iterations, we have the time bound given. ∎

In applications where the matroids are specifically given, one can often derive a better time bound, by obtaining $D_{M_1,M_2}(I')$ not from scratch, but by adapting $D_{M_1,M_2}(I)$. See also Brezovec, Cornuéjols, and Glover [1986] and Gabow and Xu [1989,1996].

41.4. Intersection of the independent set polytopes

It turns out that the intersection of the independent set polytopes of two matroids gives exactly the convex hull of the common independent sets, as was shown by Edmonds [1970b][27].

We first prove a very useful theorem, due to Edmonds [1970b], which we often will apply in this part. (A more general statement and interpretation in terms of network matrices will be given in Section 13.4.)

A family $\mathcal{C}$ of sets is called *laminar* if

(41.36) $$Y \subseteq Z \text{ or } Z \subseteq Y \text{ or } Y \cap Z = \emptyset$$

for all $Y, Z \in \mathcal{C}$.

Theorem 41.11. *Let $\mathcal{C}$ be the union of two laminar families of subsets of a set X. Let A be the $\mathcal{C} \times X$ incidence matrix of $\mathcal{C}$. Then A is totally unimodular.*

[27] Lawler [1976b] wrote that this result was announced by Edmonds 'at least as long ago as 1964'.

Proof. Let A be a counterexample with $|\mathcal{C}| + |X|$ minimal, and (secondly) with a minimal number of 1's. Then A is nonsingular and has determinant $\neq \pm 1$. Let $\mathcal{C}_1$ and $\mathcal{C}_2$ be laminar families, with union $\mathcal{C}$.

If each $\mathcal{C}_i$ consists of pairwise disjoint sets, then A is the incidence matrix of a bipartite graph, added with some unit base vectors. Hence A is totally unimodular, a contradiction.

If say $\mathcal{C}_1$ does not consist of pairwise disjoint sets, $\mathcal{C}_1$ contains a smallest nonempty set Y that is contained in some other set Z in $\mathcal{C}_1$. Choose Z smallest. Replacing Z by $Z \setminus Y$, maintains laminarity of $\mathcal{C}_1$. As this does not change the determinant of the corresponding matrix (as it amounts to subtracting row indexed Y from row indexed Z), we would have a counterexample with a smaller number of 1's, a contradiction. ∎

Let $M_1 = (S, \mathcal{I}_1)$ and $M_2 = (S, \mathcal{I}_2)$ be matroids, with rank functions r_1 and r_2. By Corollary 40.2a, the intersection $P_{\text{independent set}}(M_1) \cap P_{\text{independent set}}(M_2)$ of the independent set polytopes associated with the matroids $M_1 = (S, \mathcal{I}_1)$ and $M_2 = (S, \mathcal{I}_2)$ is determined by:

$$\text{(41.37)} \qquad \begin{array}{rll} \text{(i)} & x_s \geq 0 & \text{for } s \in S, \\ \text{(ii)} & x(U) \leq r_i(U) & \text{for } i = 1, 2 \text{ and } U \subseteq S. \end{array}$$

Trivially, this intersection contains the convex hull of the incidence vectors of common independent sets of M_1 and M_2. We shall see that these two polytopes are equal.

Basis is the following result of Edmonds [1970b], whose proof we follow (it constitutes the base of a fundamental technique developed further in several other results).

Theorem 41.12. *System* (41.37) *is box-totally dual integral.*

Proof. Choose $w \in \mathbb{Z}^S$. Consider the linear programming problem dual to maximizing $w^{\mathsf{T}} x$ over the constraints (41.37)(ii):

$$\text{(41.38)} \qquad \begin{array}{ll} \text{minimize} & \displaystyle\sum_{U \subseteq S} (y_1(U) r_1(U) + y_2(U) r_2(U)) \\ \text{where} & y_1, y_2 \in \mathbb{R}_+^{\mathcal{P}(S)}, \\ & \displaystyle\sum_{U \subseteq S} (y_1(U) + y_2(U)) \chi^U = w. \end{array}$$

Let y_1, y_2 attain this minimum, such that

$$\text{(41.39)} \qquad \sum_{U \subseteq S} (y_1(U) + y_2(U)) |U| |S \setminus U|$$

is minimized. Define

$$\text{(41.40)} \qquad \mathcal{F}_i := \{U \subseteq S \mid y_i(U) > 0\},$$

for $i = 1, 2$. We show that for $i = 1, 2$, the collection $\mathcal{F}_i$ is a chain; that is, (41.41) if $T, U \in \mathcal{F}_i$, then $T \subseteq U$ or $U \subseteq T$.

Suppose not. Choose $\alpha := \min\{y_i(T), y_i(U)\}$, and decrease $y_i(T)$ and $y_i(U)$ by α, and increase $y_i(T \cap U)$ and $y_i(T \cup U)$ by α. Since

(41.42) $$\chi^T + \chi^U = \chi^{T \cap U} + \chi^{T \cup U},$$

y_1, y_2 remains a feasible solution of (41.38); and since

(41.43) $$r_i(T) + r_i(U) \geq r_i(T \cap U) + r_i(T \cup U),$$

it remains optimum. However, sum (41.39) decreases (by Theorem 2.1), contradicting the minimality assumption. So $\mathcal{F}_1$ and $\mathcal{F}_2$ are chains.

As the constraints in (41.37)(ii) corresponding to $\mathcal{F}_1$ and $\mathcal{F}_2$ form a totally unimodular matrix (by Theorem 41.11), by Theorem 5.35 system (41.37)(ii) is box-TDI, and hence (41.37) is box-TDI. ∎

(The fact that the $\mathcal{F}_i$ can be taken to be chains also follows directly from the proof method of Theorem 40.2.)

This implies a characterization of the *common independent set polytope*

(41.44) $$P_{\text{common independent set}}(M_1, M_2)$$

of two matroids $M_1 = (S, \mathcal{I}_1)$ and $M_2 = (S, \mathcal{I}_2)$, being the convex hull of the incidence vectors of the common independent sets of M_1 and M_2:

Corollary 41.12a. $P_{\text{common independent set}}(M_1, M_2)$ *is determined by* (41.37).

Proof. Directly from Theorem 41.12, since it implies that the vertices of the polytope determined by (41.37) are integer, and hence are the incidence vectors of common independent sets. ∎

Another way of stating this is:

Corollary 41.12b.

(41.45) $$P_{\text{common independent set}}(M_1, M_2) = P_{\text{independent set}}(M_1) \cap P_{\text{independent set}}(M_2).$$

Proof. From Corollary 41.12a, using the fact that (41.37) is the union of the constraints for the independent set polytopes of M_1 and M_2, by Corollary 40.2b. ∎

The total dual integrality of (41.37) gives the following extension of the matroid intersection theorem:

Corollary 41.12c. *Let $M_1 = (S, \mathcal{I}_1)$ and $M_2 = (S, \mathcal{I}_2)$ be matroids, with rank functions r_1 and r_2, respectively, and let $w \in \mathbb{Z}_+^S$. Then the maximum value of $w(I)$ over $I \in \mathcal{I}_1 \cap \mathcal{I}_2$ is equal to the minimum value of*

(41.46) $\quad r_1(U_1) + \cdots + r_1(U_k) + r_2(T_1) + \cdots + r_2(T_l),$

where $U_1 \subseteq \cdots \subseteq U_k \subseteq S$ *and* $T_1 \subseteq \cdots \subseteq T_l \subseteq S$ *such that each element* s *of* S *occurs in precisely* $w(s)$ *sets among* $U_1, \ldots, U_k, T_1, \ldots, T_l$.

Proof. Directly from Theorem 41.12 and its proof. ∎

(Edmonds [1979] gave an algorithmic proof of this result.)

These corollaries cannot be extended to the intersection of the independent set polytopes of three matroids. Let $S = \{1, 2, 3\}$, and for $i = 1, 2, 3$, let M_i be the matroid on S with $S \setminus \{i\}$ as unique circuit. Then $P_{\text{independent set}}(M_1) \cap P_{\text{independent set}}(M_2) \cap P_{\text{independent set}}(M_3)$ contains the all-$\frac{1}{2}$ vector, while each integer vector in this intersection contains at most one 1. So the intersection is *not* the convex hull of the common independent sets.

Similar results hold for the common base polytope. For matroids M_1 and M_2, let the *common base polytope* $P_{\text{common base}}(M_1, M_2)$ be the convex hull of the incidence vectors of common bases of M_1 and M_2. Then:

Corollary 41.12d. $P_{\text{common base}}(M_1, M_2) = P_{\text{base}}(M_1) \cap P_{\text{base}}(M_2)$.

Proof. Directly from the foregoing. ∎

So the common base polytope is determined by:

(41.47)
$$\begin{array}{ll} x_s \geq 0 & \text{for } s \in S, \\ x(U) \leq r_i(U) & \text{for } i = 1, 2 \text{ and } U \subseteq S, \\ x(S) = r_i(S) & \text{for } i = 1, 2. \end{array}$$

Corollary 41.12e. *System* (41.47) *is box-TDI.*

Proof. From Theorem 41.12, with Theorem 5.25. ∎

Moreover, similar results hold for the common spanning set polytope. For matroids M_1 and M_2, let the *common spanning set polytope*, in notation $P_{\text{common spanning set}}(M_1, M_2)$, be the convex hull of the incidence vectors of common spanning sets of M_1 and M_2. Then:

Corollary 41.12f.

(41.48)
$$P_{\text{common spanning set}}(M_1, M_2) = P_{\text{spanning set}}(M_1) \cap P_{\text{spanning set}}(M_2).$$

Proof. This can be reduced to Corollary 41.12b on the common independent set polytope, by duality: x belongs to the spanning set polytope of M_i if and only if $\mathbf{1} - x$ belongs to the independent set polytope of M_i^*.

Similarly, x belongs to the common spanning set polytope of M_1 and M_2 if and only if $\mathbf{1} - x$ belongs to the common independent set polytope of M_1^* and M_2^*. ∎

So the common spanning set polytope is determined by:

$$(41.49) \qquad \begin{array}{ll} 0 \leq x_s \leq 1 & \text{for } s \in S, \\ x(U) \leq r_i(S) - r_i(S \setminus U) & \text{for } i = 1, 2 \text{ and } U \subseteq S. \end{array}$$

Corollary 41.12g. *System* (41.49) *is box-TDI.*

Proof. Again, this can be derived from Theorem 41.12, by replacing x by $\mathbf{1} - x$. ∎

Another consequence of Theorem 41.12 is:

Corollary 41.12h. *Let $M_1 = (S, \mathcal{I}_1)$ and $M_2 = (S, \mathcal{I}_2)$ be matroids and let $x \in \mathbb{R}_+^S$. Then*

$$(41.50) \qquad \begin{array}{l} \max\{z(S) \mid z \leq x, z \in P_{\text{common independent set}}(M_1, M_2)\} \\ = \min\{r(U) + x(S \setminus U) \mid U \subseteq S\}, \end{array}$$

where $r(U)$ denotes the maximum size of a common independent set contained in U.

Proof. This follows from the box-total dual integrality of (41.37), using the fact that $r(U_1 \cup U_2) \leq r_1(U_1) + r_2(U_2)$ for disjoint U_1, U_2. ∎

Cunningham [1984] showed that, if matroids $M_1 = (S, \mathcal{I}_1)$ and $M_2 = (S, \mathcal{I}_2)$ are given by independence testing oracles, one can find in strongly polynomial time for any $x \in \mathbb{Q}^S$, optimum solutions of (41.50). This will follow from the results in Section 47.4.

The result of Cunningham [1984] also implies:

Theorem 41.13. *Given matroids $M_1 = (S, \mathcal{I}_1)$ and $M_2 = (S, \mathcal{I}_2)$ by independence testing oracles, and given $x \in \mathbb{Q}^S$, one can test in strongly polynomial time if x belongs to the common independent set polytope, and if so, decompose x as a convex combination of incidence vectors of common independent sets.*

Proof. Let r_i be the rank function of M_i ($i = 1, 2$) and let $r(U) := \min\{r_1(U), r_2(U)\}$ for $i = 1, 2$. Let P be the common independent set polytope. Corollaries 40.4a and 41.12b imply that one can test in strongly polynomial time if x belongs to P.

So we can assume that x belongs to P. We decompose x as a convex combination of incidence vectors of common independent sets. Iteratively resetting x, we keep a collection $\mathcal{U}$ of subsets of S with $x(U) = r(U)$ for each $U \in \mathcal{U}$. Initially, $\mathcal{U} := \emptyset$. We describe the iteration.

Define

(41.51) $$F := \{y \in P \mid \forall s \in S : x_s = 0 \Rightarrow y_s = 0; \forall U \in \mathcal{U} : y(U) = r(U)\}.$$

So F is a face of P containing x.

Find a common independent set I with $\chi^I \in F$. This can be done by finding a common independent set $I \subseteq \mathrm{supp}(x)$ maximizing $w^{\mathsf{T}}x$, where $w := \sum_{U \in \mathcal{U}} \chi^U$. (Here $\mathrm{supp}(x)$ is the support of x; so $\mathrm{supp}(x) = \{s \in S \mid x_s > 0\}$.)

If $x = \chi^I$ we stop. Otherwise, define $u := x - \chi^I$. Let λ be the largest rational such that

(41.52) $$\chi^I + \lambda u$$

belongs to P.

We describe an inner iteration to find λ. We consider vectors z along the halfline $L = \{\chi^I + \lambda u \mid \lambda \geq 0\}$. First we let λ be the largest rational with $\chi^I + \lambda u \geq \mathbf{0}$, and set $z := \chi^I + \lambda u$.

We iteratively reset z. We check if z belongs to the common independent set polytope, and if not, we find a $U \subseteq S$ minimizing $r(U) - z(U)$ (with Corollary 40.4c). Let z' be the (unique) vector on L achieving $x(U) \leq r(U)$ with equality; that is, satisfying $z'(U) = r(U)$.

Consider any inequality $x(U') \leq r(U')$ violated by z'. Then

(41.53) $$r(U') - |U' \cap I| < r(U) - |U \cap I|.$$

This can be seen by considering the function

(41.54) $$d(y) := (r(U) - y(U)) - (r(U') - y(U')).$$

We have $d(z) \leq 0$ (since U minimizes $r(U) - z(U)$) and $d(z') > 0$ (since $z'(U) = r(U)$ and $z'(U') > r(U')$). Hence, as d is linear, $d(\chi^I) > 0$; that is, we have (41.53). This implies that resetting $z := z'$, there are at most $r(S)$ inner iterations.

Let x' be the final z found. If we apply no inner iteration, then $x'_s = 0$ for some $s \in I \subseteq \mathrm{supp}(x)$ (since we chose λ largest with $\chi^I + \lambda u \geq \mathbf{0}$). If we do at least one inner iteration, we find a U such that x' satisfies $x'(U) = r(U)$ while $|U \cap I| < r(U)$ (since x' is the unique vector on L satisfying $x'(U) = r(U)$ and since $x' \neq \chi^I$).

In the latter case, set $\mathcal{U}' := \mathcal{U} \cup \{U\}$; otherwise set $\mathcal{U}' := \mathcal{U}$. Then resetting x to x' and $\mathcal{U}$ to $\mathcal{U}'$, the dimension of F decreases (as χ^I does not belong to the new F). So the number of iterations is at most $|S|$. This shows that the method is strongly polynomial-time. ∎

41.4a. Facets of the common independent set polytope

Since the common independent set polytope of two matroids is the intersection of their independent set polytopes, each facet-inducing inequality for the intersection is facet-inducing for (at least) one of the independent set polytopes, but not necessarily conversely. Giles [1975] characterized which inequalities are facet-inducing for the common independent set polytope. If this polytope is full-dimensional, then each inequality $x_s \geq 0$ is facet-inducing. As for the other inequalities, Giles proved:

Theorem 41.14. *Let $M_1 = (S, \mathcal{I}_1)$ and $M_2 = (S, \mathcal{I}_2)$ be loopless matroids, with rank functions r_1 and r_2. For $U \subseteq S$, define $r(U) := \min\{r_1(U), r_2(U)\}$. Then, for $U \subseteq S$, the inequality*

(41.55) $\qquad x(U) \leq r(U)$

is facet-inducing for $P_{\text{common independent set}}(M_1, M_2)$ if and only if there is no partition of U into nonempty proper subsets U_1, U_2 with

(41.56) $\qquad r(U) \geq r(U_1) + r(U_2)$

and there is no proper superset U' of U with $r(U') \leq r(U)$.

Proof. By symmetry, we can assume that $r(U) = r_1(U)$.

Necessity is easy: Assume that $x(U) \leq r_1(U)$ is facet-inducing. If (41.56) would hold, then each common independent set I with $|I \cap U| = r_1(U)$ satisfies $|I \cap U_1| = r(U_1)$ (since $|I \cap U_1| = |I \cap U| - |I \cap U_2| \geq r(U) - r(U_2) \geq r(U_1)$). Hence each x in the facet determined by $x(U) \leq r_1(U)$ satisfies $x(U_1) = r(U_1)$, a contradiction. Similarly, if $r(U') \leq r_1(U)$ for some proper superset U' of U, then each common independent set I with $|I \cap U| = r_1(U)$ satisfies $|I \cap U'| = r(U')$, implying that each x in the facet determined by $x(U) \leq r_1(U)$ satisfies $x(U') = r(U')$, again a contradiction.

To see sufficiency, suppose that (41.55) satisfies the conditions, but is not facet-inducing for the common independent set polytope. This implies that the inequality $x(U) \leq r_1(U)$ is implied by other inequalities in (41.37). So there exist $\lambda_i : \mathcal{P}(S) \to \mathbb{Q}_+$ (for $i = 1, 2$) such that

(41.57)
$$\sum_{T \in \mathcal{P}(S)} (\lambda_1(T) + \lambda_2(T))\chi^T \geq \chi^U \text{ and}$$
$$\sum_{T \in \mathcal{P}(S)} (\lambda_1(T)r_1(T) + \lambda_2(T)r_2(T)) \leq r_1(U),$$

and such that $\lambda_i(U) = 0$ for $i = 1, 2$. Let D be the least common denominator of the values of the λ_i. Choose the λ_i such that D is as small as possible and (secondly) such that

(41.58)
$$D \cdot \sum_{T \subseteq S} (\lambda_1(T) + \lambda_2(T))|T|(|S \setminus T| + 1)$$

is as small as possible. For $i = 1, 2$, define

(41.59) $\qquad \mathcal{F}_i := \{T \subseteq S \mid \lambda_i(T) > 0\}.$

We claim that for $i = 1, 2$:

(41.60) $\qquad \mathcal{F}_i$ is a chain.

Suppose to the contrary that $T_1, T_2 \in \mathcal{F}_i$ satisfy $T_1 \not\subseteq T_2 \not\subseteq T_1$. Then decreasing $\lambda_i(T_1)$ and $\lambda_i(T_2)$ by $1/D$ and increasing $\lambda_i(T_1 \cap T_2)$ and $\lambda_i(T_1 \cup T_2)$ by $1/D$ maintains (41.57) but decreases (41.58). This would be a contradiction, except if $T_1 \cap T_2$ or $T_1 \cup T_2$ equals U. If one of these sets equals U and $D \geq 2$, we can reset $\lambda_i(U) := 0$, and multiply all values of λ_1 and λ_2 by $D/(D-1)$. This again maintains (41.57) but decreases the least common divisor of the denominators. So the contradiction would remain, except if $D = 1$. Then (41.57) implies $r_i(T_1) + r_i(T_2) \leq r_1(U)$. Now if $T_1 \cap T_2 = U$, then $U \subset T_1$ and

$$r(T_1) \leq r_i(T_1) \leq r_i(T_1) + r_i(T_2) \leq r_1(U), \tag{41.61}$$

contradicting the condition. If $T_1 \cup T_2 = U$, then

$$r(T_1) + r(U \setminus T_1) \leq r_i(T_1) + r_i(U \setminus T_1) \leq r_i(T_1) + r_i(T_2) \leq r_1(U), \tag{41.62}$$

again contradicting the condition.

This proves (41.60). As each $\mathcal{F}_i$ is a chain, the incidence matrix of $\mathcal{F}_1 \cup \mathcal{F}_2$ is totally unimodular (by Theorem 41.11). Therefore, there are integer-valued λ_i satisfying (41.57), with $\lambda_i(T) = 0$ for $T \notin \mathcal{F}_i$. Then we can assume that $|\mathcal{F}_i| \leq 1$ for $i = 1, 2$, since if $T, T' \in \mathcal{F}_i$ and $T \subset T'$, we can decrease $\lambda_i(T)$ by 1 without violating (41.57). If $U' \in \mathcal{F}_i$ with $U' \supset U$, then $r(U') \leq r_i(U') \leq r(U)$, contradicting the condition. So each $\mathcal{F}_i$ contains a set $U_i \not\supseteq U$, implying $r(U_1) + r(U \setminus U_1) \leq r(U_1) + r(U_2) \leq r_1(U_1) + r_2(U_2) \leq r(U)$, again contradicting the condition. ∎

This theorem can be seen to imply a variant of it, in which, instead of $r(U) := \min\{r_1(U), r_2(U)\}$, we define

$$r(U) := \max\{|I| \mid I \in \mathcal{I}_1 \cap \mathcal{I}_2\} = \min_{T \subseteq U}(r_1(T) + r_2(U \setminus T)). \tag{41.63}$$

Fonlupt and Zemirline [1983] characterized the dimension of the common base polytope of two matroids.

41.4b. Up and down hull of the common base polytope

We saw in Corollary 41.12d a characterization of the common base polytope $P_{\text{common base}}(M_1, M_2)$ of two matroids $M_1 = (S, \mathcal{I}_1)$ and $M_2 = (S, \mathcal{I}_2)$. The up hull of this polytope:

$$P^{\uparrow}_{\text{common base}}(M_1, M_2) := P_{\text{common base}}(M_1, M_2) + \mathbb{R}^S_+ \tag{41.64}$$

was characterized by Cunningham [1977] and McDiarmid [1978] as follows (proving a conjecture of Fulkerson [1971a]).

Let $M_1 = (S, \mathcal{I}_1)$ and $M_2 = (S, \mathcal{I}_2)$ be matroids having a common base. Then $P^{\uparrow}_{\text{common base}}(M_1, M_2)$ is determined by:

$$x(U) \geq r(S) - r(S \setminus U) \text{ for } U \subseteq S, \tag{41.65}$$

where $r(Z) :=$ the maximum size of a common independent set contained in Z. (A weaker version of this was proved by Edmonds and Giles [1977].)

For a proof we refer to Section 46.7a, where it is also shown that (41.65) is TDI (Gröflin and Hoffman [1981]). (Frank and Tardos [1984a] derived this, with a direct algorithmic construction, from the total dual integrality of (41.47).)

Note that by the matroid intersection theorem, the inequalities (41.65) are equivalent to:

$$x(U) \geq k - r_1(A) - r_2(B) \text{ for each partition } U, A, B \text{ of } S, \tag{41.66}$$

where r_1 and r_2 are the rank functions of M_1 and M_2 respectively, and where k is the size of a common base. This implies that if we add $x \leq \mathbf{1}$ to (41.66) we obtain the convex hull of the subsets of S that contain a common base.

Similarly, the down hull of the common base polytope:

$$(41.67) \qquad P^{\downarrow}_{\text{common base}}(M_1, M_2) := P_{\text{common base}}(M_1, M_2) - \mathbb{R}^S_+,$$

is determined by

$$(41.68) \qquad x(U) \leq r_1(S \setminus A) + r_2(S \setminus B) - k \text{ for each partition } U, A, B \text{ of } S.$$

This can be derived from the description of the up hull of the common base polytope, since

$$(41.69) \qquad P^{\downarrow}_{\text{common base}}(M_1, M_2) = \mathbf{1} - P^{\uparrow}_{\text{common base}}(M_1^*, M_2^*)$$

(where $\mathbf{1}$ stands for the all-one vector in $\mathbb{R}^S$).

This implies that the convex hull of the incidence vectors of the subsets of common bases is determined by $x \geq \mathbf{0}$ and (41.68).

Cunningham [1984] gave a strongly polynomial-time algorithm to test if a vector belongs to $P^{\uparrow}_{\text{common base}}(M_1, M_2)$, or to $P^{\downarrow}_{\text{common base}}(M_1, M_2)$, using only independence testing oracles for M_1 and M_2.

41.5. Further results and notes

41.5a. Menger's theorem for matroids

Tutte [1965b] showed a special case of the matroid intersection theorem, namely when both M_1 and M_2 are minors of one matroid. Specialized to graphic matroids, it gives the vertex-disjoint, undirected version of Menger's theorem.

Let $M = (E, \mathcal{I})$ be a matroid, with rank function r, and let U and W be disjoint subsets of E. Then the maximum size of a common independent set in $M/U \setminus W$ and $M/W \setminus U$ is equal to the minimum value of

$$(41.70) \qquad r(X) - r(U) + r(E \setminus X) - r(W)$$

taken over sets X with $U \subseteq X \subseteq E \setminus W$. This is the special case of the matroid intersection theorem for the matroids $M/U \setminus W$ and $M/W \setminus U$, since for $Y \subseteq E \setminus (U \cup W)$ one has

$$(41.71) \qquad r_{M/U \setminus W}(Y) = r(Y \cup U) - r(U),$$

and similarly for $M/W \setminus U$.

To see that this implies the vertex-disjoint, undirected version of Menger's theorem, let $G = (V, E)$ be a graph and let S and T be disjoint nonempty subsets of V. We show that the above theorem implies that the maximum number of disjoint $S - T$ paths in G is equal to the minimum number of vertices intersecting each $S - T$ path.

To this end, we can assume that G is connected, and that E contains subsets U and W such that (S, U) and (T, W) are trees. (Adding appropriate edges does not modify the result to be proved.)

Let $M := M(G)$ be the cycle matroid of G. Define $R := V \setminus (S \cup T)$. Then

(41.72) the maximum number of disjoint $S-T$ paths is at least the maximum size of a common independent set I of $M/U \setminus W$ and $M/W \setminus U$, minus $|R|$.

(In fact, there is equality.)

To prove (41.72), let I be a maximum-size common independent in $M/U \setminus W$ and $M/W \setminus U$. So I is a forest. Consider any component K of I. Since I is independent in M/U, K intersects S in at most one vertex. Similarly, K intersects T in at most one vertex. Let p be the number of components K intersecting both S and T. By deleting p edges we obtain a forest I' such that no component of I' intersects both S and T. So $|I'| \leq |R|$ (since I' remains a forest after contracting (in the graphical sense) $S \cup T$ to one vertex). Hence $p = |I| - |I'| \geq |I| - |R|$. So we have (41.72).

On the other hand,

(41.73) the minimum size of a set of vertices intersecting each $S-T$ path is at most the minimum value of (41.70), minus $|R|$.

(Again, we have in fact equality.)

To prove (41.73), let X attain the minimum value of (41.70). So $U \subseteq X \subseteq E \setminus W$. Let K be the component of (V, X) containing S and let L be the component of $(V, E \setminus X)$ containing T. We choose X with $|K \cup L|$ maximized.

Then $K \cup L = V$. For suppose not. Then, as G is connected, there is an edge e of G leaving $K \cup L$. By symmetry, we can assume that $e \in X$. Let K' be the component of (V, X) containing e. So $K' \neq K$ and $E[K'] \cap U = \emptyset$. Resetting X by $X \setminus E[K']$, $r(X)$ decreases by $|K'| - 1$, while $r(E \setminus X)$ increases by at most $|K'| - 1$. So the new X again attains the minimum in (41.70), while $K \cup L$ increases. This contradicts our maximality assumption.

So $K \cup L = V$. Hence $K \cap L$ intersects each $S-T$ path (since $S \subseteq K$ and $T \subseteq L$, and there is no edge connecting $K \setminus L$ and $L \setminus K$). Moreover

(41.74) $$\begin{aligned}|K \cap L| &= |K| + |L| - |V| \leq (r(X) + 1) + (r(E \setminus X) + 1) - |V| \\ &= r(X) + r(E \setminus X) - |V| + 2 = r(X) + r(E \setminus X) - r(U) - r(W) - |R|.\end{aligned}$$

So we have (41.73).

Since the maximum number of disjoint $S-T$ paths is trivially not more than the minimum number of vertices intersecting all $S-T$ paths, we thus obtain Menger's theorem (and also equality in (41.72) and (41.73)).

(Tomizawa [1976a] gave an algorithm for Menger's theorem for matroids.)

41.5b. Exchange properties

Kundu and Lawler [1973] showed the following extension of the exchange property of bipartite graphs given in Theorem 16.8. Let $M_1 = (S, \mathcal{I}_1)$ and $M_2 = (S, \mathcal{I}_2)$ be matroids, with span functions span_1 and span_2. Then

(41.75) For any $I_1, I_2 \in \mathcal{I}_1 \cap \mathcal{I}_2$ there exists an $I \in \mathcal{I}_1 \cap \mathcal{I}_2$ with $I_1 \subseteq \mathrm{span}_1(I)$ and $I_2 \subseteq \mathrm{span}_2(I)$.

(Theorem 16.8 is equivalent to the case where the M_i are partition matroids.)

To prove (41.75), choose $I \in \mathcal{I}_1 \cap \mathcal{I}_2$ with $I_1 \subseteq \mathrm{span}_1(I)$ and $|I \cap I_2|$ maximized. Suppose that $I_2 \not\subseteq \mathrm{span}_2(I)$. Choose $s \in I_2 \setminus \mathrm{span}_2(I)$ with $I \cup \{s\} \in \mathcal{I}_2$. By the maximality of $|I \cap I_2|$ we know that $I \cup \{s\} \notin \mathcal{I}_1$. So M_1 has a circuit C

contained in $I \cup \{s\}$. Since $I_2 \in \mathcal{I}_1$ we know that $C \not\subseteq I_2$. Choose $t \in C \setminus I_2$. Then for $I' := I - t + s$ we have $I' \in \mathcal{I}_1 \cap \mathcal{I}_2$, while $\mathrm{span}_1(I') = \mathrm{span}_1(I)$. Since $|I' \cap I_2| > |I \cap I_2|$ this contradicts the maximality assumption.

A second exchange property was shown by Davies [1976]:

(41.76) Two matroids M_1 and M_2 have bases B_1 and B_2 (respectively) with $|B_1 \cap B_2| = k$ if and only if M_1 has bases X_1 and Y_1 and M_2 has bases X_2 and Y_2 with $|X_1 \cap X_2| \leq k$ and $|Y_1 \cap Y_2| \geq k$.

To see this, we may assume that $X_2 = Y_2$, since if $|X_1 \cap Y_2| \leq k$ we can reset $X_2 := Y_2$, and if $|X_1 \cap Y_2| > k$ we can reset $Y_1 := X_1$ and exchange indices.

By (39.33)(ii), there exists a series of bases $Z_0, \ldots, Z_t$ of M_1 such that $Z_0 = X_1$, $Z_t = Y_1$, and $|Z_{i-1} \triangle Z_i| = 2$ for $i = 1, \ldots, t$. Hence

(41.77) $$\big||Z_{i-1} \cap X_2| - |Z_i \cap X_2|\big| \leq 1$$

for $i = 1, \ldots, t$. Since $|Z_0 \cap X_2| \leq k$ and $|Z_t \cap X_2| \geq k$, we know $|Z_i \cap X_2| = k$ for some i. This proves (41.76).

41.5c. Jump systems

A framework that includes both matroid intersection and maximum-size matching was introduced by Bouchet and Cunningham [1995]. For $x, y \in \mathbb{Z}^n$, let $[x, y]$ be the set of vectors $z \in \mathbb{Z}^n$ with $\|x - y\|_1 = \|x - z\|_1 + \|z - y\|_1$. So $[x, y]$ consists of all integer vectors z in the box $x \wedge y \leq z \leq x \vee y$.

Call a vector z a *step from x to y* if $z \in [x, y]$ and $\|z - x\|_1 = 1$. A *jump system* is a finite subset J of $\mathbb{Z}^n$ satisfying the following axiom:

(41.78) if $x, y \in J$ and z is a step from x to y, then $z \in J$ or J contains a step from z to y.

Trivially, for any jump system J and any $x, y \in \mathbb{Z}^n$, the intersection $J \cap [x, y]$ is again a jump system. Moreover, being a jump system is maintained under translations by an integer vector and by reflections in a coordinate hyperplane. Bouchet and Cunningham [1995] showed that the sum of jump systems is again a jump system (attributing the proof below to A. Sebő):

Theorem 41.15. *If J_1 and J_2 are jump systems in $\mathbb{Z}^n$, then $J_1 + J_2$ is a jump system.*

Proof. For $x, y \in J_1 + J_2$ we prove (41.78) by induction on the minimum value of

(41.79) $$\|y' - x'\|_1 + \|y'' - x''\|_1,$$

where $x', y' \in J_1$, $x'', y'' \in J_2$, $x' + x'' = x$, and $y' + y'' = y$.

Let z be a step from x to y. By reflection and permutation of coordinates, we can assume that $z = x + \chi^1$. So $x_1 < y_1$. Hence, by symmetry of J_1 and J_2, we can assume that $x'_1 < y'_1$. Next, by reflection, we can assume that $x' \leq y'$.

Now $x' + \chi^1$ is a step from x' to y'. If $x' + \chi^1 \in J_1$, then $z = x' + \chi^1 + x'' \in J_1 + J_2$, and we have (41.78). So we can assume that $x' + \chi^1 \notin J_1$. Hence, by (41.78) applied to J_1, there is an $i \in \{1, \ldots, n\}$ with $\tilde{x}' := x' + \chi^1 + \chi^i \in J_1$ and $\tilde{x}' \leq y'$.

So $z + \chi^i = \tilde{x}' + x'' \in J_1 + J_2$. If $z + \chi^i \in [x, y]$, we have (41.78). If $z + \chi^i \notin [x, y]$, then as $z \in [x, y]$, we have $z_i = y_i$. So z is a step from $z + \chi^i$ to y. Also,

$\|y' - \tilde{x}'\|_1 = \|y' - x'\|_1 - 2$. Hence, by our induction hypothesis applied to $z + \chi^i$ and y, we have (41.78). ∎

As Bouchet and Cunningham [1995] observed, this theorem implies that the following two constructions give jump systems $J \subseteq \mathbb{Z}^V$.

For any matroid $M = (S, \mathcal{I})$, the set $\{\chi^B \mid B \text{ base of } M\}$ is a jump system in $\mathbb{Z}^S$, as follows directly from the axioms (39.33). With Theorem 41.15, this implies that for matroids $M_1 = (S, \mathcal{I}_1)$ and $M_2 = (S, \mathcal{I}_2)$, the set

(41.80) $$J := \{\chi^{B_1} - \chi^{B_2} \mid B_i \text{ base of } M_i \ (i = 1, 2)\}$$

is a jump system.

Let $G = (V, E)$ be an undirected graph and let

(41.81) $$J := \{\deg_F \mid F \subseteq E\} \subseteq \mathbb{Z}^V;$$

that is, J is the collection of degree sequences of spanning subgraphs of G. Again, J is a jump system. This follows from Theorem 41.15, since for each edge $e = uv$ the set $\{\mathbf{0}, \chi^{\{u,v\}}\}$ is trivially a jump system in $\mathbb{Z}^V$ and since J is the sum of these jump systems.

Bouchet and Cunningham [1995] showed that the following greedy approach finds, for any $w \in \mathbb{R}^n$, a vector $x \in J$ maximizing $w^\mathsf{T} x$. By reflecting, we can assume that $w \geq \mathbf{0}$. We can also assume that $w_1 \geq w_2 \geq \cdots \geq w_n$. Let $J_0 := J$, and for $i = 1, \ldots, n$, let J_i be the set of vectors x in J_{i-1} maximizing x_i over J_{i-1}. Trivially, J_n consists of one vector, y say. Then:

Theorem 41.16. *y maximizes $w^\mathsf{T} x$ over J.*

Proof. It suffices to show that the maximum value of $w^\mathsf{T} x$ over J_1 is the same as over J (since applying this to the jump systems $J_1, \ldots, J_n$ gives the theorem). Let the maximum over J be attained by x and over J_1 by y. Suppose $w^\mathsf{T} y < w^\mathsf{T} x$. So $x \notin J_1$, and hence $x_1 < y_1$. We choose x, y such that $y_1 - x_1$ is minimal. Let $z := x + \chi^1$. So z is a step from x to y.

Then $w^\mathsf{T} z = w^\mathsf{T} x + w_1 \geq w^\mathsf{T} x$. Hence $z \notin J$, since otherwise we can replace x by z, contradicting the minimality of $y_1 - x_1$. So, by (41.78), J contains a step u from z to y. So $u = z \pm \chi^i$ for some $i \in \{1, \ldots, n\}$. Then

(41.82) $$w^\mathsf{T} u = w^\mathsf{T} z \pm w_i \geq w^\mathsf{T} z - w_i = w^\mathsf{T} x + w_1 - w_i \geq w^\mathsf{T} x.$$

So we can replace x by u, again contradicting the minimality of $y_1 - x_1$ (as $u_1 > x_1$). ∎

Lovász [1997] gave a min-max relation for the minimum l_1-distance of an integer vector to a jump system of special type. It can be considered as a common generalization of the matroid intersection theorem (Theorem 41.1) and the Tutte-Berge formula (Theorem 24.1).

For a survey, see Cunningham [2002].

41.5d. Further notes

A special case of the weighted matroid intersection algorithm (where one matroid is a partition matroid) was studied by Brezovec, Cornuéjols, and Glover [1988].

Data structures for on-line updating of matroid intersection solutions were given by Frederickson and Srinivas [1984,1987], and a randomized parallel algorithm for linear matroid intersection by Narayanan, Saran, and Vazirani [1992,1994].

An extension of matroid intersection to 'supermatroid' intersection was given by Tardos [1990]. Fujishige [1977a] gave a primal approach to weighted matroid intersection, and Shigeno and Iwata [1995] a dual approximation approach. Camerini and Maffioli [1975,1978] studied 3-matroid intersection problems.

Chapter 42

Matroid union

Matroid union is closely related to matroid intersection, and most of the basic matroid union results follow from basic matroid intersection results, and vice versa. But matroid union also gives a shift in focus and offers a number of specific algorithmic questions.

42.1. Matroid union theorem

The matroid union theorem will be derived from the following basic result given by Nash-Williams [1967], suggested by earlier unpublished work of J. Edmonds[28]:

Theorem 42.1. *Let $M' = (S', \mathcal{I}')$ be a matroid, with rank function r', and let $f : S' \to S$. Define*

(42.1) $\qquad \mathcal{I} := \{f(I') \mid I' \in \mathcal{I}'\}$

(where $f(I') := \{f(s) \mid s \in I'\}$). Then $M = (S, \mathcal{I})$ is a matroid, with rank function r given by

(42.2) $\qquad r(U) = \min_{T \subseteq U}(|U \setminus T| + r'(f^{-1}(T)))$

for $U \subseteq S$.

Proof. Trivially, $\mathcal{I}$ is nonempty and closed under taking subsets. To see condition (39.1)(ii), let $I, J \in \mathcal{I}$ with $|I| < |J|$. Choose $I', J' \in \mathcal{I}'$ with $f(I') = I$, $f(J') = J$, $|I'| = |I|$, $|J'| = |J|$, and $|I' \cap J'|$ as large as possible. As M' is a matroid, $I' + j \in \mathcal{I}'$ for some $j \in J' \setminus I'$. If $f(j) \in f(I')$, say $f(j) = f(i)$ for $i \in I'$, replacing I' by $I' - i + j$ would increase $|I' \cap J'|$, contradicting our assumption. So $f(j) \in J \setminus I$ and $f(I') + f(j) = f(I' + j) \in \mathcal{I}$. This proves (39.1)(ii), and hence M is a matroid.

The rank $r(U)$ of a subset U of S is equal to the maximum size of a common independent set in M' and the partition matroid $N = (S', \mathcal{J})$ induced by the family $(f^{-1}(s) \mid s \in U)$. By the matroid intersection theorem (Theorem 41.1), this is equal to the right-hand side of (42.2). ∎

[28] as mentioned in the footnote on page 20 of Pym and Perfect [1970] (quoted in Section 42.6f below).

(In his paper, Nash-Williams suggested a direct proof, by decomposing f as a product of 'elementary' functions in which only two elements are merged. Welsh [1970] observed that the rank formula (42.2) also follows directly from Rado's theorem (Corollary 41.1c) of Rado [1942].)

Theorem 42.1 implies the following result, formulated explicitly by Edmonds [1968] (and for all M_i equal by Nash-Williams [1967]).

Let $M_1 = (S_1, \mathcal{I}_1), \ldots, M_k = (S_k, \mathcal{I}_k)$ be matroids. Define the *union* of these matroids as $M_1 \vee \cdots \vee M_k = (S_1 \cup \cdots \cup S_k, \mathcal{I}_1 \vee \cdots \vee \mathcal{I}_k)$, where

(42.3) $$\mathcal{I}_1 \vee \cdots \vee \mathcal{I}_k := \{I_1 \cup \ldots \cup I_k \mid I_1 \in \mathcal{I}_1, \ldots, I_k \in \mathcal{I}_k\}.$$

Corollary 42.1a (matroid union theorem)**.** *Let $M_1 = (S_1, \mathcal{I}_1), \ldots, M_k = (S_k, \mathcal{I}_k)$ be matroids, with rank functions $r_1, \ldots, r_k$, respectively. Then $M_1 \vee \cdots \vee M_k$ is a matroid again, with rank function r given by:*

(42.4) $$r(U) = \min_{T \subseteq U}(|U \setminus T| + r_1(T \cap S_1) + \cdots + r_k(T \cap S_k)).$$

for $U \subseteq S_1 \cup \cdots \cup S_k$.

Proof. To see that $M_1 \vee \cdots \vee M_k$ is a matroid, let for each i, $M_i' = (S_i', \mathcal{I}_i')$ be a copy of M_i with $S_1', \ldots, S_k'$ disjoint. Then trivially $M_1' \vee \cdots \vee M_k'$ is a matroid. Now define $f : S_1' \cup \cdots \cup S_k' \to S_1 \cup \cdots \cup S_k$ by, for $i = 1, \ldots, k$ and $s \in S_i'$: $f(s)$ is the original of s in S_i. Then the matroid obtained in Theorem 42.1 is equal to $M_1 \vee \cdots \vee M_k$, proving that the latter indeed is a matroid, and (42.4) follows from (42.2). ∎

Conversely, the matroid intersection theorem may be derived from the matroid union theorem (as was shown by Edmonds [1970b]): the maximum size of a common independent set in two matroids M_1 and M_2, is equal to the maximum size of an independent set in the union $M_1 \vee M_2^*$, minus the rank of M_2^*.

Application of the matroid union theorem to a number of copies of the same matroid gives the following results. First:

Corollary 42.1b. *Let $M = (S, \mathcal{I})$ be a matroid, with rank function r, and let $k \in \mathbb{Z}_+$. Then the maximum size of the union of k independent sets is equal to*

(42.5) $$\min_{U \subseteq S}(|S \setminus U| + k \cdot r(U)).$$

Proof. This follows by applying Corollary 42.1a to $M_1 = \cdots = M_k = M$. ∎

This implies that the minimum number of independent sets (or bases) needed to cover the underlying set is described by the following result of Edmonds [1965c][29]:

[29] This result was also given, without proof, by Rado [1966], saying that the argument of Horn [1955] for linear matroids can be extended to arbitrary matroids. The result con-

Corollary 42.1c (matroid base covering theorem)**.** *Let $M = (S, \mathcal{I})$ be a matroid, with rank function r, and let $k \in \mathbb{Z}_+$. Then S can be covered by k independent sets if and only if*

$$k \cdot r(U) \geq |U| \tag{42.6}$$

for each $U \subseteq S$.

Proof. M can be covered by k independent sets if and only if there is a union of k independent sets of size $|S|$. By Corollary 42.1b, this is the case if and only if

$$\min_{U \subseteq S}(|S \setminus U| + k \cdot r(U)) \geq |S|, \tag{42.7}$$

that is, if and only if $k \cdot r(U) \geq |U|$ for each subset U of S. ∎

One similarly has for the maximum number of disjoint bases in a matroid (Edmonds [1965a]):

Corollary 42.1d (matroid base packing theorem)**.** *Let $M = (S, \mathcal{I})$ be a matroid, with rank function r, and let $k \in \mathbb{Z}_+$. Then there exist k disjoint bases if and only if*

$$k \cdot (r(S) - r(U)) \leq |S \setminus U| \tag{42.8}$$

for each $U \subseteq S$.

Proof. M has k disjoint bases if and only if the maximum size of the union of k independent sets is equal to $k \cdot r(S)$. By Corollary 42.1b, this is the case if and only if

$$\min_{U \subseteq S}(|S \setminus U| + k \cdot r(U)) \geq k \cdot r(S), \tag{42.9}$$

that is, if and only if $|S \setminus U| \geq k \cdot (r(S) - r(U))$ for each subset U of S. ∎

The more general forms of Corollaries 42.1c and 42.1d, with different matroids, were shown by Edmonds and Fulkerson [1965].

42.1a. Applications of the matroid union theorem

We describe a number of applications of the matroid union theorem. Further applications will follow in Chapter 51 on packing and covering of trees and forests.

Transversal matroids. Let $\mathcal{X} = (X_1, \ldots, X_n)$ be a family of subsets of a finite set S, and define for each $i = 1, \ldots, n$ a matroid M on S by: Y is independent in M_i if and only if $Y \subseteq X_i$ and $|Y| \leq 1$. Now the union $M_1 \vee \cdots \vee M_n$ is the same

firms a question of Rado [1962a,1962b] (in fact, the result also follows by an elementary construction from Rado's theorem (Corollary 41.1c) given in Rado [1942]).

as the transversal matroid induced by $\mathcal{X}$, so in this way one can prove again that transversal matroids indeed are matroids.

Disjoint transversals. Let $\mathcal{X} = (X_1, \ldots, X_n)$ be a family of subsets of a finite set S. Then $\mathcal{X}$ has k disjoint transversals if and only if

$$(42.10) \qquad |\bigcup_{i \in I} X_i| \geq k \cdot |I|$$

for each $I \subseteq \{1, \ldots, n\}$. This easy consequence of Hall's marriage theorem (cf. Theorem 22.10) can also be derived by applying the matroid base packing theorem to the transversal matroid induced by $\mathcal{X}$, using (39.19).

Similarly, it can be derived from the matroid base covering theorem that S can be partitioned into k partial transversals of $\mathcal{X}$ if and only if

$$(42.11) \qquad k(n - |I|) \geq |S \setminus \bigcup_{i \in I} X_i|$$

for each $I \subseteq \{1, \ldots, n\}$ (cf. Theorem 22.12).

Vector spaces. A finite subset S of a vector space can be covered by k linearly independent sets if and only if

$$(42.12) \qquad |U| \leq k \cdot \mathrm{rank}(U) \text{ for each } U \subseteq S.$$

This conjecture of K.F. Roth and R. Rado was shown by Horn [1955][30]. It is the special case of the matroid base covering theorem for linear matroids (see also Section 42.1b below).

As a similar consequence of the matroid base packing theorem one has that the n-dimensional vector space S over the field $GF(q)$ contains $k := \lfloor (q^n - 1)/n \rfloor$ disjoint bases. Indeed, for each $U \subseteq S$ one has $k(n - r(U)) \leq q^n - |U|$, as $|U| \leq q^{r(U)}$.

An exchange property of bases. The matroid union theorem also implies the following stronger exchange property of bases of a matroid (stronger than given in the 'axioms' in Theorem 39.6). In any matroid $M = (S, \mathcal{I})$,

(42.13) for any two bases B_1 and B_2 and for any partition of B_1 into X_1 and Y_1, there is a partition of B_2 into X_2 and Y_2 such that both $X_1 \cup Y_2$ and $X_2 \cup Y_1$ are bases.

This property was conjectured by G.-C. Rota, and proved by Brylawski [1973], Greene [1973], and Woodall [1974a] — we follow the proof of McDiarmid [1975a].

Consider the matroids $M_1 := M/Y_1$ and $M_2 := M/X_1$. Note that M_1 has rank $|X_1|$ and that M_2 has rank $|Y_1|$. We must show that B_2 is the union of an independent set X_2 of M_1 and an independent set Y_2 of M_2. By the submodularity of the rank functions ((39.38)(ii)) we have for each $T \subseteq B_2$:

[30] Horn [1955] thanked Rado 'for improvements in the setting out of the argument'. The result was also published, in the same journal, by Rado [1962a]. This paper does not mention Horn's paper. The proof by Rado [1962a] is the same as that of Horn [1955] and uses the same notation. But Rado [1966] said that the theorem was first proved by Horn [1955].

(42.14)
$$\begin{aligned}&|B_2 \setminus T| + r_{M_1}(T \setminus Y_1) + r_{M_2}(T \setminus X_1)\\ &= |B_2 \setminus T| + r(T \cup Y_1) - |Y_1| + r(T \cup X_1) - |X_1|\\ &\geq r(T) + r(T \cup Y_1 \cup X_1) - |T| = |B_2|.\end{aligned}$$

Hence, by the matroid union theorem (Corollary 42.1a), we have the required result.

Repeated application of this exchange phenomenon implies the following stronger property, given by Greene and Magnanti [1975]:

(42.15) for any two bases B_1 and B_2 and any partition of B_1 into $X_1, \ldots, X_k$, there is a partition of B_2 into $Y_1, \ldots, Y_k$ such that $(B_1 \setminus X_i) \cup Y_i$ is a base, for each $i = 1, \ldots, k$.

This extends Corollary 39.12a, which is the special case where each X_i is a singleton.

42.1b. Horn's proof

The proof of Horn [1955] of the matroid base covering theorem for linear matroids directly extends to general matroids (as was observed by Rado [1966]):

Consider a counterexample to the matroid base covering theorem (Corollary 42.1c) with smallest $|S|$. For subsets $S_1, \ldots, S_n$ of S, define inductively:

(42.16)
$$[S_1, \ldots, S_n] := \begin{cases} S & \text{if } n = 0,\\ \text{span}([S_1, \ldots, S_{n-1}] \cap S_n) & \text{if } n \geq 1.\end{cases}$$

By the minimality of $|S|$, we know that for each $s \in S$, $S \setminus \{s\}$ can be partitioned into k independent sets $I_1, \ldots, I_k$. We first show:

(42.17) for each $s \in S$ and $I_1, \ldots, I_k$ partitioning $S \setminus \{s\}$, there exist $j_1, \ldots, j_n \in \{1, \ldots, k\}$ with $s \notin [I_{j_1}, \ldots, I_{j_n}]$.

Indeed, choose $j_1, \ldots, j_n \in \{1, \ldots, k\}$ with the rank of $[I_{j_1}, \ldots, I_{j_n}]$ as small as possible. Define $A := [I_{j_1}, \ldots, I_{j_n}]$. By the minimality of the rank of A, we have $r(A \cap I_j) = r(A)$ for each $j = 1, \ldots, k$. Hence, by (42.6),

(42.18)
$$|A| \leq k \cdot r(A) = \sum_{j=1}^{k} r(A \cap I_j) \leq \sum_{j=1}^{k} |A \cap I_j| = |A \setminus \{s\}|.$$

So $s \notin A$, proving (42.17).

Now choose s, $I_1, \ldots, I_k$, and $j_1, \ldots, j_n$ as in (42.17) with n as small as possible. For $t = 0, \ldots, n$, define

(42.19)
$$B_t := [I_{j_1}, \ldots, I_{j_t}].$$

As we have a counterexample, we know that $s \in \text{span}(I_{j_n})$ (otherwise we can add s to I_{j_n}). Let C be the circuit in $I_{j_n} \cup \{s\}$. As $s \notin B_n = \text{span}(B_{n-1} \cap I_{j_n})$, we know that $C \setminus \{s\}$ is not contained in B_{n-1} (otherwise $C \setminus \{s\} \subseteq B_{n-1} \cap I_{j_n}$, and hence $s \in \text{span}(B_{n-1} \cap I_{j_n})$). So we can choose $z \in C \setminus \{s\}$ with $z \notin B_{n-1}$.

Define $I'_{j_n} := I_{j_n} - z + s$ and $I'_j := I_j$ for $j \neq j_n$. Then $I'_1, \ldots, I'_k$ are independent sets partitioning $S \setminus \{z\}$. Define, for $t = 0, \ldots, n$:

(42.20)
$$B'_t := [I'_{j_1}, \ldots, I'_{j_t}].$$

By the minimality of n we know that $z \in B'_{n-1}$. Since $z \notin B_{n-1}$, we have $B'_{n-1} \not\subseteq B_{n-1}$. Choose the smallest $q \leq n-1$ with $B'_q \not\subseteq B_q$. Then $q \geq 1$ and $B'_{q-1} \subseteq B_{q-1}$. By the minimality of n we know that $s \in B_q$ (as $q < n$). So

(42.21)
$$B'_q = \text{span}(B'_{q-1} \cap I'_{j_q}) \subseteq \text{span}((B_{q-1} \cap I_{j_q}) \cup \{s\}) \subseteq \text{span}(B_q) = B_q,$$

a contradiction.

42.2. Polyhedral applications

The matroid base packing and covering theorems imply (in fact, are equivalent to) the following polyhedral result:

Corollary 42.1e. *For any matroid, the independent set polytope, the base polytope, and the spanning set polytope have the integer decomposition property.*

Proof. Let $M = (S, \mathcal{I})$ be a matroid. Choose $k \in \mathbb{Z}_+$ and an integer vector $x \in k \cdot P_{\text{independent set}}(M)$. Replace each element s of S by x_s parallel elements, thus obtaining the matroid $N = (T, \mathcal{J})$ say. Now for each $U \subseteq T$, one has $k \cdot r_N(U) \geq |U|$, since if W denotes the set of elements s in S such that U intersects the parallel class of s, then

(42.22) $$r_N(U) = r_M(W) \geq x(W)/k \geq |U|/k,$$

since x/k belongs to $P_{\text{independent set}}(M)$. So by the matroid base covering theorem (Corollary 42.1c), T can be partitioned into k independent sets of N. Hence x is the sum of k incidence vectors of independent sets of M.

To see that the base polytope has the integer decomposition property, let $x \in k \cdot P_{\text{base}}(M)$. By the above, x is the sum of the incidence vectors of k independent sets. As $x(S) = k \cdot r(S)$, each of these independent sets is a base.

One similarly derives from the matroid base packing theorem (Corollary 42.1d) that the spanning set polytope has the integer decomposition property. ∎

The matroid base packing and covering theorems imply generalizations to the capacitated case, by splitting elements into parallel elements. For the matroid base covering theorem this gives:

Theorem 42.2. *Let $M = (S, \mathcal{I})$ be a matroid, with rank function r, and let $c : S \to \mathbb{Z}_+$. Then the minimum value of $\sum_{I \in \mathcal{I}} \lambda_I$, where $\lambda : \mathcal{I} \to \mathbb{Z}_+$ satisfies*

(42.23) $$\sum_{I \in \mathcal{I}} \lambda_I \chi^I = c,$$

is equal to the maximum value of

(42.24) $$\left\lceil \frac{c(U)}{r(U)} \right\rceil$$

taken over $U \subseteq S$ with $r(U) \geq 1$.

Proof. Directly from the matroid base covering theorem (Corollary 42.1c), by splitting each $s \in S$ into $c(s)$ parallel elements. ∎

In other words, the system defining the antiblocking polyhedron of the independent set polytope:

$$\text{(42.25)}\qquad \begin{array}{ll} x_s \geq 0 & \text{for } s \in S, \\ x(I) \leq 1 & \text{for } I \in \mathcal{I}, \end{array}$$

has the integer rounding property (the optimum integer solution to the dual of maximizing $c^{\mathsf{T}}x$ over (42.25) has value equal to the upper integer part of the value of the optimum (fractional) solution, for any integer objective function c).

Similarly, the matroid base packing theorem gives:

Theorem 42.3. *Let $M = (S, \mathcal{I})$ be a matroid, with rank function r, and let $c : S \to \mathbb{Z}_+$. Let $\mathcal{B}$ be the collection of bases of M. Then the maximum value of $\sum_{B \in \mathcal{B}} \lambda_B$, where $\lambda : \mathcal{B} \to \mathbb{Z}_+$ satisfies*

$$\text{(42.26)}\qquad \sum_{B \in \mathcal{B}} \lambda_B \chi^B \leq c,$$

is equal to the minimum value of

$$\text{(42.27)}\qquad \left\lfloor \frac{c(S \setminus U)}{r(S) - r(U)} \right\rfloor$$

taken over $U \subseteq S$ with $r(S) - r(U) \geq 1$.

Proof. Directly from the matroid base packing theorem (Corollary 42.1d), by splitting each $s \in S$ into $c(s)$ parallel elements. ∎

In other words, the system defining the blocking polyhedron of the base polytope:

$$\text{(42.28)}\qquad \begin{array}{ll} x_s \geq 0 & \text{for } s \in S, \\ x(B) \geq 1 & \text{for } B \in \mathcal{B}, \end{array}$$

has the integer rounding property.

De Pina and Soares [2000] showed that, in Theorem 42.3, the number of bases B with $\lambda_B > 0$ can be restricted to at most $|S| + r$, where r is the rank of M. This strengthens a result of Cook, Fonlupt, and Schrijver [1986].

42.3. Matroid union algorithm

A polynomial-time algorithm for partitioning a matroid in as few independent sets as possible may be derived from the matroid intersection algorithm, with the construction given in the proof of Theorem 42.1. A direct algorithm was given by Edmonds [1968]. We give the algorithm described by Knuth [1973] and Greene and Magnanti [1975], which is similar to the algorithm described in Section 41.2 for cardinality matroid intersection.

Let $M_1 = (S, \mathcal{I}_1), \ldots, (S, \mathcal{I}_k)$ be matroids. Let $I_i \in \mathcal{I}_i$, for $i = 1, \ldots, k$, with $I_i \cap I_j = \emptyset$ if $i \neq j$. Let D be the union of the graphs $D_{M_i}(I_i)$ as defined in Section 39.9.

For each i, let F_i be the set of elements $s \notin I_i$ with $I_i \cup \{s\} \in \mathcal{I}_i$. Define $I := I_1 \cup \cdots \cup I_k$, $F := F_1 \cup \cdots \cup F_k$, and $\mathcal{I} := \mathcal{I}_1 \vee \cdots \vee \mathcal{I}_k$.

Theorem 42.4. *For any $s \in S \setminus I$ one has: $I \cup \{s\} \in \mathcal{I} \iff D$ has an $F - s$ path.*

Proof. To see necessity, suppose that D has no $F - s$ path. Let T be the set of elements of S that can reach s in D. So $s \in T$, $T \cap F = \emptyset$, and no arc of D enters T. Then $r_i(T) = |I_i \cap T|$ for each $i = 1, \ldots, k$. Otherwise, there exists a $t \in T \setminus I_i$ with $(I_i \cap T) \cup \{t\} \in \mathcal{I}_i$. Since $t \notin F$, $I_i \cup \{t\} \notin \mathcal{I}_i$. So there is a $u \in I_i \setminus T$ with $I_i - u + t \in \mathcal{I}_i$. But then (u, t) is an arc of D entering T, a contradiction.

So $r_i(T) = |I_i \cap T|$ for each i. Hence $r_1(T) + \cdots + r_k(T) = |I \cap T|$. As $s \in T \setminus I$, this implies $(I \cap T) \cup \{s\} \notin \mathcal{I}$, and so $I \cup \{s\} \notin \mathcal{I}$.

To see sufficiency, let $P = (s_0, s_1, \ldots, s_p)$ be a shortest $F - s$ path in D. We can assume by symmetry that $s_0 \in F_1$; so $s_0 \notin I_1$ and $I_1 \cup \{s_0\} \in \mathcal{I}_1$. Since P is a shortest path, for each $i = 1, \ldots, k$, the set N_i of edges (s_{j-1}, s_j) with $j = 1, \ldots, p$ and $s_{j-1} \in I_i$, forms a unique perfect matching in $D_{M_i}(I_i)$ on the set S_i covered by N_i. So by Theorem 39.13, $I_i \triangle S_i$ belongs to $\mathcal{I}_i$ for each i. Moreover, by Corollary 39.13a, $(I_1 \triangle S_1) \cup \{s_0\} \in \mathcal{I}_1$. So $I \cup \{s\} \in \mathcal{I}$. ∎

This implies that a maximum-size set in $\mathcal{I}_1 \vee \cdots \vee \mathcal{I}_k$ can be found in polynomial time (by greedily growing an independent set in $M_1 \vee \cdots \vee M_k$). Similarly, we can find with the greedy algorithm a maximum-weight set in $\mathcal{I}_1 \vee \cdots \vee \mathcal{I}_k$.

In particular, we can test if a given set is independent in $M_1 \vee \cdots \vee M_k$. Cunningham [1986] gave an $O((n^{3/2} + k)mQ + n^{1/2}km)$ algorithm to find a maximum-size set in $\mathcal{I}_1 \vee \cdots \vee \mathcal{I}_k$, where n is the maximum size of a set in $\mathcal{I}_1 \vee \cdots \vee \mathcal{I}_k$, m is the size of the underlying set, and Q is the time needed to test if a given set belongs to $\mathcal{I}_j$ for any given j.

These methods (including the reduction to matroid intersection) also imply:

Theorem 42.5. *Given a matroid $M = (S, \mathcal{I})$ by an independence testing oracle, we can find a maximum number of disjoint bases, and a minimum number of independent sets covering S, in polynomial time.*

Proof. See above. ∎

42.4. The capacitated case: fractional packing and covering of bases

The complexity of the capacitated and fractional cases of the above packing and covering problems can be studied with the help of the strong polynomial-time solvability of the *most violated inequality problem* for a matroid $M = (S, \mathcal{I})$, with rank function r:

(42.29) given: a vector $x \in \mathbb{Q}_+^S$;
find: a subset U of S minimizing $r(U) - x(U)$.

The strong polynomial-time solvability of this problem was shown in Corollary 40.4c, and is a result of Cunningham [1984].

If x belongs to $P_{\text{independent set}}(M)$, we can decompose x as a convex combination of incidence vectors of independent sets. This decomposition can be found in strongly polynomial time, by Corollary 40.4a.

We now consider the problem of finding a maximum fractional packing of bases subject to a given capacity function, and its dual, finding a minimum fractional covering by independent sets of a demand function.

With a method given by Picard and Queyranne [1982a] and Padberg and Wolsey [1984] one finds:

Theorem 42.6. *Given a matroid $M = (S, \mathcal{I})$ by an independence testing oracle and given $y \in \mathbb{Q}_+^S$, we can find the minimum value of λ such that $y \in \lambda \cdot P_{\text{independent set}}(M)$ in strongly polynomial time.*

Proof. Let r be the rank function of M. We can assume that y does not belong to the independent set polytope. Let L be the line through 0 and y. We iteratively reset y as follows. By Corollary 40.4c, we can find a subset U of S minimizing $r(U) - y(U)$. Let y' be the vector on L with $y'(U) = r(U)$.

Now, for any $U' \subseteq S$, if y' violates $x(U') \leq r(U')$, then $r(U') < r(U)$, since the function $d(x) := (r(U) - x(U)) - (r(U') - x(U'))$ is nonpositive at y and positive at y', implying that it is positive at 0 (as d is linear in x).

We reset $y := y'$ and iterate, until y belongs to $P_{\text{independent set}}(M)$. So after at most $r(S)$ iterations the process terminates, with a y on the boundary of $P_{\text{independent set}}(M)$. Comparing the final y with the original y gives the required λ. ∎

Theorem 42.6 implies an algorithm for capacitated fractional covering by independent sets:

Corollary 42.6a. *Given a matroid $M = (S, \mathcal{I})$ by an independence testing oracle and given $y \in \mathbb{Q}_+^S$, we can find independent sets $I_1, \ldots, I_k$ and rationals $\lambda_1, \ldots, \lambda_k \geq 0$ such that*

(42.30) $$y = \lambda_1 \chi^{I_1} + \cdots + \lambda_k \chi^{I_k}$$

with $\lambda_1 + \cdots + \lambda_k$ minimal, in strongly polynomial time.

Proof. Without loss of generality, $y \neq \mathbf{0}$. By Theorem 51.7, we can find the minimum value of λ such that y belongs to $\lambda \cdot P_{\text{independent set}}(M)$. By Corollary 40.4a, we can decompose $\frac{1}{\lambda} \cdot y$ as a convex combination of incidence vectors of independent sets. This gives the required decomposition of y. ∎

One similarly shows for the spanning set polytope:

Theorem 42.7. *Given a matroid $M = (S, \mathcal{I})$ by an independence testing oracle and given $y \in \mathbb{Q}_+^S$, we can find the maximum value of λ such that $y \in \lambda \cdot P_{\text{spanning set}}(M)$, in strongly polynomial time.*

Proof. Let r be the rank function of M. By Corollary 40.2f, the spanning set polytope of M is determined by the constraints $\mathbf{0} \leq x \leq \mathbf{1}$ and

$$r(U) - x(U) \geq r(S) - x(S) \text{ for } U \subseteq S. \tag{42.31}$$

We can assume that $y \notin P_{\text{spanning set}}(M)$ and that the support of y is a spanning set. Let L be the line through 0 and y. We iteratively reset y as follows.

Find a $U \subseteq S$ minimizing $r(U) - y(U)$ (this can be done in strongly polynomial time, by Corollary 40.4c). If y does not belong to the spanning set polytope, we know that y violates the constraint $r(U) - x(U) \geq r(S) - x(S)$. Let y' be the vector on L satisfying $r(U) - y'(U) = r(S) - y'(S)$.

Now for any $U' \subseteq S$, if y' violates $r(U') - x(U') \geq r(S) - x(S)$, then $r(U') > r(U)$, since the function $d(x) := (r(U) - x(U)) - (r(U') - x(U'))$ is nonpositive at y and positive at y', implying that it is negative at 0 (as d is linear in x).

We reset $y := y'$ and iterate, until y belongs to $P_{\text{spanning set}}(M)$. So after at most $r(S)$ iterations the process terminates, in which case y is on the boundary of $P_{\text{spanning set}}(M)$. Comparing the final y with the original y gives the required λ. ∎

In turn, this gives an algorithm for capacitated fractional base packing:

Corollary 42.7a. *Given a matroid $M = (S, \mathcal{I})$ by an independence testing oracle and given $y \in \mathbb{Q}_+^S$, we can find bases $B_1, \ldots, B_k$ and rationals $\lambda_1, \ldots, \lambda_k \geq 0$ such that*

$$y \geq \lambda_1 \chi^{B_1} + \cdots + \lambda_k \chi^{B_k} \tag{42.32}$$

with $\lambda_1 + \cdots + \lambda_k$ maximal, in strongly polynomial time.

Proof. By Theorem 42.7, we can find the maximum value of λ such that y belongs to $\lambda \cdot P_{\text{spanning set}}(M)$. If $\lambda = 0$, we take $k = 0$. If $\lambda > 0$, by Corollary 40.4b we can decompose $\frac{1}{\lambda} \cdot y$ as a convex combination of incidence vectors of spanning sets. This gives the required decomposition of y. ∎

42.5. The capacitated case: integer packing and covering of bases

It is not difficult to derive integer versions of the above algorithms, but they are not strongly polynomial-time, as we round numbers in it. In fact, an integer packing or covering cannot be found in strongly polynomial time, as it would imply a strongly polynomial-time algorithm for testing if an integer k is even (which algorithm does not exist[31]): Let M be the 2-uniform matroid on 3 elements and let $k \in \mathbb{Z}_+$. Then k is even if and only if M has $\frac{3}{2}k$ bases containing each element of M at most k times.

Polynomial-time algorithms follow directly from the fractional versions with the help of the matroid base packing and covering theorems.

Theorem 42.8. *Given a matroid $M = (S, \mathcal{I})$ by an independence testing oracle and given $y \in \mathbb{Z}_+^S$, we can find independent sets $I_1, \ldots, I_t$ and integers $\lambda_1, \ldots, \lambda_t \geq 0$ such that*

$$y = \lambda_1 \chi^{I_1} + \cdots + \lambda_t \chi^{I_t} \tag{42.33}$$

with $\lambda_1 + \cdots + \lambda_t$ minimal, in polynomial time.

Proof. First find $I_1, \ldots, I_k$ and $\lambda_1, \ldots, \lambda_k$ as in Corollary 42.6a. We can assume that $k \leq |S|$ (by Carathéodory's theorem, applying Gaussian elimination). Let

$$y' := \sum_{i=1}^{k} (\lambda_i - \lfloor \lambda_i \rfloor) \chi^{I_i} = y - \sum_{i=1}^{k} \lfloor \lambda_i \rfloor \chi^{I_i}. \tag{42.34}$$

So y' is integer.

Replace each $s \in S$ by $y'(s)$ parallel elements, making matroid $M' = (S', \mathcal{I}')$. By Theorem 42.5, we can find a minimum number of independent sets partitioning S', in polynomial time (as $y'(s) \leq |S|$ for each $s \in S$). This gives independent sets $I_{k+1}, \ldots, I_t$ of M.

Setting $\lambda_i := 1$ for $i = k+1, \ldots, t$, we show that this gives a solution of our problem. Trivially, (42.33) is satisfied (with λ_i replaced by $\lfloor \lambda_i \rfloor$). By the matroid base covering theorem applied to M' (as (42.34) gives a fractional decomposition of S' into independent sets),

$$t - k \leq \left\lceil \sum_{i=1}^{k} (\lambda_i - \lfloor \lambda_i \rfloor) \right\rceil. \tag{42.35}$$

Therefore,

$$\sum_{i=1}^{t} \lfloor \lambda_i \rfloor = (t - k) + \sum_{i=1}^{k} \lfloor \lambda_i \rfloor \leq \left\lceil \sum_{i=1}^{k} \lambda_i \right\rceil, \tag{42.36}$$

[31] For any strongly polynomial-time algorithm with one integer k as input, there is a number L and a rational function $q : \mathbb{Z} \to \mathbb{Q}$ such that if $k > L$, then the output equals $q(k)$. (This can be proved by induction on the number of steps of the algorithm, which is a fixed number as the input consists of only one number.) However, there do not exist a rational function q and number L such that for $k > L$, $q(k) = 0$ if k is even, and $q(k) = 1$ if k is odd.

proving that the decomposition is optimum (cf. Theorem 42.2). ∎

One similarly shows for packing bases:

Theorem 42.9. *Given a matroid $M = (S, \mathcal{I})$ by an independence testing oracle and given $y \in \mathbb{Z}_+^S$, we can find bases $B_1, \ldots, B_t$ and integers $\lambda_1, \ldots, \lambda_t \geq 0$ such that*

$$y \geq \lambda_1 \chi^{B_1} + \cdots + \lambda_t \chi^{B_t} \tag{42.37}$$

with $\lambda_1 + \cdots + \lambda_t$ maximal, in polynomial time.

Proof. First find bases $B_1, \ldots, B_k$ and $\lambda_1, \ldots, \lambda_k$ as in Corollary 42.7a. Again we can assume that $k \leq |S|$. Let

$$y' := \left\lceil \sum_{i=1}^{k} (\lambda_i - \lfloor \lambda_i \rfloor) \chi^{B_i} \right\rceil. \tag{42.38}$$

Replace each $s \in S$ by $y'(s)$ parallel elements, making matroid M'. By Theorem 42.5, we can find a maximum number of disjoint bases in M' in polynomial time (as $y'(s) \leq |S|$ for each $s \in S$). This gives bases $B_{k+1}, \ldots, B_t$ in M.

Setting $\lambda_i := 1$ for $i = k+1, \ldots, t$, we show that this gives a solution of our problem. Trivially, (42.37) is satisfied (with λ_i replaced by $\lfloor \lambda_i \rfloor$). Again, now by the matroid base packing theorem applied to M', using (42.38),

$$t - k \geq \left\lfloor \sum_{i=1}^{k} (\lambda_i - \lfloor \lambda_i \rfloor) \right\rfloor. \tag{42.39}$$

Therefore,

$$\sum_{i=1}^{t} \lfloor \lambda_i \rfloor = (t - k) + \sum_{i=1}^{k} \lfloor \lambda_i \rfloor \geq \left\lfloor \sum_{i=1}^{k} \lambda_i \right\rfloor, \tag{42.40}$$

proving that the decomposition is optimum (cf. Theorem 42.3). ∎

De Pina and Soares [2000] showed that, in this theorem we can make the additional condition that $t \leq |S| + r$, where r is the rank of M.

42.6. Further results and notes

42.6a. Induction of matroids

An application of matroid intersection and union is the following 'induction of a matroid through a directed graph', discovered by Perfect [1969b] (for bipartite graphs) and Brualdi [1971c]. In fact, it forms a generalization of the basic Theorem 42.1.

Let $D = (V, A)$ be a directed graph, let $U, W \subseteq V$, and let $M = (U, \mathcal{I})$ be a matroid. Let $\mathcal{J}$ be the collection of subsets Y of W such that there exists an $X \in \mathcal{I}$ with X linked to Y. (Set X is *linked to* Y if $|X| = |Y|$ and D has $|X|$ disjoint $X-Y$ paths.)

Then:

(42.41) $\quad N = (W, \mathcal{J})$ is a matroid.

To show that N is a matroid, we can assume that U and W are disjoint. (Otherwise, add a new vertex w' and new arc (w, w') for each $w \in W$.) Let L be the gammoid induced by $D, U, U \cup W$. Then $N = (M \vee L)/U$. Indeed, since U is independent in L and hence in $M \vee L$, a subset Y of W is independent in $(M \vee L)/U$ if and only if $Y \cup U$ is independent in $M \vee L$. This is easily seen to be equivalent to: $Y \in \mathcal{J}$. So N is a matroid.

The rank function r_N of N can be described by (for $Y \subseteq W$):

(42.42) $\quad r_N(Y) = \min\{r_M(X) + |Z| \mid X \subseteq U, Z \subseteq V, Z$ intersects each $U \setminus X - Y$ path$\}$.

This can be derived from the matroid union theorem, but also (and simpler) from the matroid intersection theorem, as follows. Let K be the gammoid induced by D^{-1}, Y, U, where D^{-1} arises from D by reversing the orientations of all arcs. Then $r_N(Y)$ is equal to the maximum size of a common independent set in M and K. So, by the matroid intersection theorem (Theorem 41.1),

(42.43) $$r_N(Y) = \min_{X \subseteq U}(r_M(X) + r_K(U \setminus X)),$$

which by Menger's theorem is equal to the right-hand side of (42.42).

Applying the matroid intersection theorem again gives the following result of Brualdi [1971e] (generalizing Brualdi [1970a]).

Let $D = (V, A)$ be a directed graph, let $U, W \subseteq V$, and let $M = (U, \mathcal{I})$ and $M' = (W, \mathcal{I}')$ be matroids. Then the maximum size of an independent set in M that is linked to an independent set in M' is equal to the minimum value of

(42.44) $$r_M(X) + |Z| + r_{M'}(Y),$$

where $X \subseteq U$, $Y \subseteq W$, and $Z \subseteq V$, such that Z intersects each $U \setminus X - W \setminus Y$ path. (This follows directly by considering the maximum size of a common independent set in M' and N as defined above.)

Related results are given by McDiarmid [1975b] and Woodall [1975]. These results are generalized in Schrijver [1979c]. For an algorithm, see Fujishige [1977b].

42.6b. List-colouring

Seymour [1998] showed the following matroid list-colouring theorem (cf. Section 20.9c):

Theorem 42.10. *Let $M = (S, \mathcal{I})$ be a matroid such that S can be partitioned into k independent sets, and let $m \in \mathbb{Z}_+$. For each $s \in S$, let $L_s \subseteq \{1, \ldots, m\}$ be a set of size k. Then S can be partitioned into independent sets $I_1, \ldots, I_m$ such that for each $j = 1, \ldots, m$: if $s \in I_j$, then $j \in L_s$.*

Proof. For each $j = 1, \dots, m$, let $U_j := \{s \in S \mid j \in L_s\}$. We need to prove that for all j, there exists an independent set $I_j \subseteq U_j$ such that $S = I_1 \cup \dots \cup I_n$.

Since S can be partitioned into k independent sets, we know that $|X| \leq k \cdot r_M(X)$ for each $X \subseteq S$. Hence, for each $T \subseteq S$,

$$\text{(42.45)} \qquad \sum_{j=1}^{m} r_M(U_j \cap T) \geq \sum_{j=1}^{m} \frac{1}{k}|U_j \cap T| = |T|,$$

since each $s \in T$ belongs to k of the U_j. So by the matroid union theorem (Corollary 42.1a), applied to the matroids $M|U_j$, the independent sets I_j as required exist. ∎

42.6c. Strongly base orderable matroids

In general it is not true that given two matroids $M_1 = (S, \mathcal{I}_1)$ and $M_2 = (S, \mathcal{I}_2)$ such that S can be partitioned into k independent sets of M_1, and also into k independent sets of M_2, then S can be partitioned into k common independent sets of M_1 and M_2. This could yield a 'matroid union intersection theorem'. However, taking for M_1 is the cycle matroid of K_4 and for M_2 the matroid with independent sets all sets of pairwise intersecting edges of K_4 (which is a partition matroid), shows that the statement is false for $k = 2$.

But the assertion is true if both M_1 and M_2 belong to the class of so-called strongly base orderable matroids, introduced by Brualdi [1970b]. A matroid $M = (S, \mathcal{I})$ is called *strongly base orderable* if for each two bases B_1, B_2 of M there exists a bijection $\pi : B_1 \to B_2$ such that for each subset X of B_1 the set $\pi(X) \cup (B_1 \setminus X)$ is a base again.

One easily checks that for such π, the function $\pi|B_1 \cap B_2$ is the identity map. It is also straightforward to check that if M is strongly base orderable, then also the dual of M and any contraction of M is strongly base orderable, and hence also any restriction, and therefore any minor is strongly base orderable. Moreover, Brualdi [1970b] showed:

Theorem 42.11. *Any truncation of a strongly base orderable matroid is strongly base orderable again.*

Proof. Let $M = (S, \mathcal{I})$ be a strongly base orderable matroid, with rank function r, and let $k := r(S) - 1$. It suffices to show that the k-truncation of M is strongly base orderable. Let I and J be independent sets of size k, and restrict M to $I \cup J$. If $r(I \cup J) = k$, we are done, since then I and J are bases of the strongly base orderable matroid $M|I \cup J$. So suppose $r(I \cup J) = r(S) = k + 1$, and let $i \in I \setminus J$ and $j \in J \setminus I$ be such that $I \cup \{j\}$ and $J \cup \{i\}$ are bases of M. As M is strongly base orderable, there exists a bijection $\pi : I \cup \{j\} \to J \cup \{i\}$ with the prescribed exchange property. So $\pi(j) = j$ and $\pi(i) = i$. Define $\pi' : I \to J$ by $\pi'(s) := \pi(s)$ if $s \neq i$, and $\pi'(i) = j$. We show that this bijection is as required. To prove this, choose $X \subseteq I$. We must show that $\pi'(X) \cup (I \setminus X)$ is independent.

If $i \notin X$, then $\pi'(X) = \pi(X)$, hence $\pi'(X) \cup (I \setminus X)$ is independent, since

$$\text{(42.46)} \qquad \pi'(X) \cup (I \setminus X) = \pi(X) \cup (I \setminus X) \subseteq \pi(X) \cup ((I \cup \{j\}) \setminus X)$$

and the last set is independent.

If $i \in X$, then $\pi'(X) = \pi(X \setminus \{i\}) \cup \{j\}$, hence $\pi'(X) \cup (I \setminus X)$ is independent, since

(42.47) $$\pi'(X) \cup (I \setminus X) = \pi(X \setminus \{i\}) \cup \{j\} \cup (I \setminus X) = \pi(X \setminus \{i\}) \cup ((I \cup \{j\}) \setminus X) \subseteq \pi(X \setminus \{i\}) \cup ((I \cup \{j\}) \setminus (X \setminus \{i\}))$$

and the last set is independent. ∎

One also easily checks that strong base orderability is closed under making parallel extensions. (Given a matroid $M = (S, \mathcal{I})$ a *parallel extension* in $s \in S$ is obtained by extending S with some new element s', and $\mathcal{I}$ with $\{(I \setminus \{s\}) \cup \{s'\} \mid s \in I \in \mathcal{I}\}$.)

Since transversal matroids are strongly base orderable, also gammoids are strongly base orderable (Brualdi [1971c]):

Theorem 42.12. *Each gammoid is strongly base orderable.*

Proof. Since strong base orderability is closed under taking contractions and since each gammoid is a contraction of a transversal matroid (by Corollary 39.5a), it suffices to show that any transversal matroid is strongly base orderable.

Let M be the transversal matroid induced by a family $\mathcal{X} = (X_1, \ldots, X_m)$ of subsets of a set S. We may assume that $\mathcal{X}$ has a transversal (cf. (39.18)). Consider two transversals $T_1 = \{x_1, \ldots, x_m\}$ and $T_2 = \{y_1, \ldots, y_m\}$ of $\mathcal{X}$, where $x_i, y_i \in X_i$ for $i = 1, \ldots, m$.

Consider the bipartite graph on $\{1, \ldots, m\} \cup S$ with edges all pairs $\{i, s\}$ with $i \in \{1, \ldots, m\}$ and $s \in X_i$ (assuming without loss of generality that $\{1, \ldots, m\} \cap S = \emptyset$). Then $M_1 := \{\{i, x_i\} \mid i = 1, \ldots, m\}$ and $M_2 := \{\{i, y_i\} \mid i = 1, \ldots, m\}$ are matchings in G. Define $\pi : T_1 \to T_2$ as follows. If $s \in T_1 \cap T_2$, define $\pi(s) := s$. If $s \in T_1 \setminus T_2$, let $\pi(s)$ be the (other) end of the path in $M_1 \cup M_2$ starting at s. This defines a bijection as required. ∎

Brualdi [1971c] showed more generally that strong base orderability is maintained under induction of matroids through a directed graph, as described in Section 42.6a. However, not every strongly base orderable matroid is a gammoid (cf. Oxley [1992] p. 411).

Davies and McDiarmid [1976] (cf. McDiarmid [1976]) showed the following.

Theorem 42.13. *Let $M_1 = (S, \mathcal{I}_1)$ and $M_2 = (S, \mathcal{I}_2)$ be strongly base orderable matroids, let $k \in \mathbb{Z}_+$, and suppose that S can be split into k independent sets of M_1, and also into k independent sets of M_2. Then S can be split into k common independent sets of M_1 and M_2.*

Proof. In order to prove this, let $\mathcal{X} = (X_1, \ldots, X_k)$ and $\mathcal{Y} = (Y_1, \ldots, Y_k)$ be partitions of S into independent sets of M_1 and M_2, respectively, with

(42.48) $$\sum_{i=1}^{k} |X_i \cap Y_i|$$

as large as possible. If this sum is equal to $|S|$ we are done, so suppose that this sum is less than $|S|$. Hence there are i and j with $X_i \cap Y_j \neq \emptyset$ and $i \neq j$. Extend X_i and X_j to bases C_i and C_j of M_1. Similarly, extend Y_i and Y_j to bases D_i and D_j of M_2. Since M_1 and M_2 are strongly base orderable, there exist bijections

$\pi_1 : C_i \to C_j$ and $\pi_2 : D_i \to D_j$ with the exchange property. So $p_1(s) = s$ for each $s \in C_i \cap C_j$ and $p_2(s) = s$ for each $s \in D_i \cap D_j$.

Let G be the bipartite graph with vertex set $C_i \cup C_j \cup D_i \cup D_j$, and edges the pairs $\{s, \pi_1(s)\}$ with s in $C_i \setminus C_j$ and the pairs $\{s, \pi_2(s)\}$ with s in $D_i \setminus D_j$. Split the vertex set into colour classes S and T, say. Define

$$\text{(42.49)} \qquad \begin{aligned} &X'_i := S \cap (X_i \cup X_j),\ X'_j := T \cap (X_i \cup X_j),\\ &Y'_i := S \cap (Y_i \cup Y_j),\ Y'_j := T \cap (Y_i \cup Y_j). \end{aligned}$$

So $X'_i \cap Y'_j = \emptyset$ and $X'_j \cap Y'_i = \emptyset$. Moreover, X'_i and X'_j are independent in M_1, since, by the exchange property of π, $S \cap (C_i \cup C_j)$ and $T \cup (C_i \cup C_j)$ are independent in M_1. Similarly, Y'_i and Y'_j are independent in M_2.

So replacing the classes X_i and X_j of $\mathcal{X}$ by X'_i and X'_j, and the classes Y_i and Y_j of $\mathcal{Y}$ by Y'_i and Y'_j yields partitions as required. However, since $X'_i \cap Y'_j = \emptyset$ and $X'_j \cap Y'_i = \emptyset$, we have

$$\text{(42.50)} \qquad |X'_i \cap Y'_i| + |X'_j \cap Y'_j| > |X_i \cap Y_i| + |X_j \cap Y_j|,$$

contradicting the maximality of (42.48). ∎

The proof also shows that the required partition can be found in polynomial time, provided that there is a polynomial-time algorithm to find the exchange bijection π. (This is the case for transversal matroids induced by a given family of sets.)

By the matroid base covering theorem (Corollary 42.1c), Theorem 42.13 is equivalent to:

Corollary 42.13a. *Let $M_1 = (S, \mathcal{I}_1)$ and $M_2 = (S, \mathcal{I}_2)$ be loopless, strongly base orderable matroids, with rank functions r_1 and r_2. Then the minimum number of common independent sets needed to cover S, is equal to*

$$\text{(42.51)} \qquad \max\{\left\lceil \frac{|U|}{r_i(U)} \right\rceil \mid \emptyset \neq U \subseteq S, i = 1, 2\}.$$

Proof. Directly from Theorem 42.13 with the matroid base covering theorem. ∎

Applying Corollary 42.13a to transversal matroids gives Corollary 23.9a. Similarly, it follows from Theorem 42.13 that:

Corollary 42.13b. *Let $M_1 = (S, \mathcal{I}_1)$ and $M_2 = (S, \mathcal{I}_2)$ be strongly base orderable matroids, with rank functions r_1 and r_2, satisfying $r_1(S) = r_2(S)$. Then M_1 and M_2 have k disjoint common bases if and only if*

$$\text{(42.52)} \qquad |S \setminus (T \cup U)| \geq k(r_1(S) - r_1(T) - r_2(U))$$

for all $T, U \subseteq S$.

Proof. Indeed, from Theorem 42.13 we have that M_1 and M_2 have k disjoint common bases if and only if the matroids $M_1 \vee \cdots \vee M_1$ and $M_2 \vee \cdots \vee M_2$ (k-fold unions) have a common independent set of size $k \cdot r_1(S)$. By the matroid union and intersection theorems, this last is equivalent to the condition stated in the present corollary. ∎

By truncating M_1 and M_2 one has similar results if we replace 'common bases' by 'common independent sets of size t'. Application to transversal matroids yields Corollary 23.9d.

Another consequence of Theorem 42.13 is:

Corollary 42.13c. *Let $M_1 = (S, \mathcal{I}_1)$ and $M_2 = (S, \mathcal{I}_2)$ be strongly base orderable matroids. Then M_1 and M_2 have k disjoint common spanning sets if and only if both M_1 and M_2 have k disjoint bases.*

Proof. This can be deduced as follows. Let N_i arise from the dual matroid of M_i by replacing each element s of S by $k-1$ parallel elements (for $i = 1, 2$). So N_1 and N_2 are strongly base orderable again, with an underlying ground set of size $(k-1)|S|$. Now M_1 and M_2 have k disjoint (common) spanning sets, if and only if N_1 and N_2 have k (common) independent sets covering the underlying set. This directly implies the present corollary. ∎

Applying Corollary 42.13c to transversal matroids gives Theorem 23.11.

Corollary 42.13d. *Let $M_1 = (S, \mathcal{I}_1)$ and $M_2 = (S, \mathcal{I}_2)$ be strongly base orderable matroids, with rank functions r_1 and r_2, satisfying $r_1(S) = r_2(S)$. Then S can be covered by k common bases of M_1 and M_2 if and only if*

(42.53) $$k(r_1(T) + r_2(U) - r_1(S)) \geq |T \cap U|$$

for all $T, U \subseteq S$.

Proof. Condition (42.53) is equivalent to:

(42.54) $$(k-1)|S \setminus (T \cup U)| \geq k(r_1^*(S) - r_1^*(T) - r_2^*(U))$$

for all $T, U \subseteq S$. Let N_1 and N_2 be the matroids defined in the proof of Corollary 42.13c. By Corollary 42.13b, condition (42.54) implies that N_1 and N_2 contain k disjoint common bases. So M_1^* and M_2^* have k common bases covering each element at most $k-1$ times. Hence M_1 and M_2 have k common bases covering S. ∎

Applying Corollary 42.13d to transversal matroids gives Theorem 23.12.

42.6d. Blocking and antiblocking polyhedra

We next investigate the blocking and antiblocking polyhedra corresponding to intersections of independent set polytopes of two matroids. Let $M_1 = (S, \mathcal{I}_1)$ and $M_2 = (S, \mathcal{I}_2)$ be loopless matroids, with rank functions r_1 and r_2 respectively, and independent set polytopes P_1 and P_2 respectively. So $P_1 \cap P_2$ is the convex hull of the incidence vectors of common independent sets. Hence its antiblocking polyhedron $A(P_1 \cap P_2)$ is determined by the linear inequalities

(42.55) $$\begin{aligned} x_s &\geq 0 && (s \in S), \\ x(I) &\leq 1 && (I \in \mathcal{I}_1 \cap \mathcal{I}_2). \end{aligned}$$

Since $P_1 \cap P_2$ is determined by the linear inequalities (41.37), $A(P_1 \cap P_2)$ consists of all vectors $x \geq \mathbf{0}$ for which there exists a $y \geq x$ which is a convex combination of vectors

$$\frac{1}{r_i(U)}\chi^U \tag{42.56}$$

where U is a nonempty subset of S and $i = 1, 2$. Then $A(P_1 \cap P_2)$ gives rise to the following linear programming duality equation, for $c : S \to \mathbb{R}_+$:

$$\begin{aligned}\max\{c^{\mathsf T}x \mid x \in A(P_1 \cap P_2)\} &= \max\{\frac{c(U)}{r_i(U)} \mid \emptyset \neq U \subseteq S; i = 1, 2\}\\ &= \min\{\sum_{I\in\mathcal{I}_1\cap\mathcal{I}_2} y(I) \mid y \in \mathbb{R}_+^{\mathcal{I}_1\cap\mathcal{I}_2}, \sum_{I\in\mathcal{I}_1\cap\mathcal{I}_2} y(I)\chi^I \geq c\}.\end{aligned} \tag{42.57}$$

For integer c, an integer optimum solution y need not exist (for instance, if $|S| = 3$, $r_i(U) := \min\{|U|, 2\}$, and $c = \mathbf{1}$). That is, system (42.55) need not be totally dual integral. In fact, it generally does not have the integer rounding property. That is, it is not true, for each pair of matroids, that the minimum in (42.57) with y restricted to be integer:

$$\min\{\sum_{I\in\mathcal{I}_1\cap\mathcal{I}_2} y(I) \mid y \in \mathbb{Z}_+^{\mathcal{I}_1\cap\mathcal{I}_2}, \sum_{I\in\mathcal{I}_1\cap\mathcal{I}_2} y(I)\chi^I \geq c\}, \tag{42.58}$$

is equal to the upper integer part of the common value of (42.57). For instance, take for M_1 the cycle matroid of K_4, and for M_2 the matroid with independent sets all sets of pairwise intersecting edges in K_4, and let $c = \mathbf{1}$; then the common value in (42.57) is 2, while (42.58) is equal to 3. However, Corollary 42.13a implies that if M_1 and M_2 are strongly base orderable matroids, then (42.58) is equal to the upper integer part of (42.57). That is, for strongly base orderable matroids, system (42.57) has the integer rounding property.

Similar results hold if we consider the blocker $B(Q_1 \cap Q_2)$ of the intersection of the spanning set polytopes Q_1 and Q_2 of M_1 and M_2. In particular, Corollary 42.13c implies that the system

$$\begin{aligned} x_s &\geq 0 && (s \in S),\\ x(U) &\geq 1 && (U \text{ common spanning set of } M_1 \text{ and } M_2)\end{aligned} \tag{42.59}$$

has the integer rounding property, if M_1 and M_2 are strongly base orderable.

Moreover, Corollaries 42.13b and 42.13d imply that the systems

$$\begin{aligned} x_s &\geq 0 && (s \in S),\\ x(B) &\geq 1 && (B \text{ common base of } M_1 \text{ and } M_2)\end{aligned} \tag{42.60}$$

and

$$\begin{aligned} x_s &\geq 0 && (s \in S),\\ x(B) &\leq 1 && (B \text{ common base of } M_1 \text{ and } M_2)\end{aligned} \tag{42.61}$$

have the integer rounding property, if M_1 and M_2 are strongly base orderable. Here the results of Section 41.4b are used: to prove that (42.60) has the integer rounding property, let $w \in \mathbb{Z}_+^S$. Let Q be the polytope determined by (42.60), let $r(U)$ be the maximum size of a common independent set contained in U, and let $\mathcal{B}$ denote the collection of common bases. Then

$$\begin{aligned} &\lceil \min\{w^{\mathsf T}x \mid x \in Q\}\rceil\\ &= \min\{\lceil \frac{w(U)}{r(S) - r(S\setminus U)}\rceil \mid U \subseteq S, r(S) > r(S\setminus U)\}\\ &= \max\{\sum_{B\in\mathcal{B}} y_B \mid y \in \mathbb{Z}_+^{\mathcal{B}}, \sum_{B\in\mathcal{B}} y_B\chi^B \leq w\}.\end{aligned} \tag{42.62}$$

The first equality holds as the vertices of Q are given by the vectors

$$(42.63) \qquad \frac{1}{r(S) - r(S \setminus U)} \chi^U,$$

since Q is the blocking polyhedron of the common base polytope (cf. Section 41.4b). The second equality follows from Corollary 42.13b, using the fact that strong base orderability is maintained under adding parallel elements.

Related results on integer decomposition of the intersection of the independent set polytopes of two strongly base orderable matroids can be found in McDiarmid [1983].

42.6e. Further notes

Krogdahl [1976] observed that the following, general problem is solvable in polynomial time, by reduction to matroid intersection: given matroids $(S, \mathcal{I}_1), \ldots, (S, \mathcal{I}_k)$, weight functions $w_1, \ldots, w_k \in \mathbb{R}^S$, and $l \leq k$, find the maximum value of $w_1(I_1) + \cdots + w_k(I_k)$, where $I_1 \in \mathcal{I}_1, \ldots, I_k \in \mathcal{I}_k$, with $I_1, \ldots, I_l$ disjoint and $I_{l+1}, \ldots, I_k$ disjoint, and with $I_1 \cup \ldots \cup I_l = I_{l+1} \cup \ldots \cup I_k$.

With matroid union, several new classes of matroids can be constructed. One of them is formed by the *bicircular matroids*, which are the union of the cycle matroid $M(G)$ of a graph $G = (V, E)$ and the matroid on E in which $F \subseteq E$ is independent if and only if $|F| \leq 1$. The independent sets of this matroid are the edge sets containing at most one circuit.

A randomized parallel algorithm for linear matroid union was given by Narayanan, Saran, and Vazirani [1992,1994]. For matroid base packing algorithms, see Knuth [1973] and Karger [1993,1998].

42.6f. Historical notes on matroid union

As the matroid base covering theorem can be derived by an elementary construction from Rado's theorem (proved by Rado [1942]), it is surprising that, for a long time, it had remained an open question, posed by Rado himself.

In fact, it was Horn [1955] who showed that a set X of vectors is the union of k linearly independent sets of vectors if and only if each finite subset Y of X has rank at least $|Y|/k$. He mentioned that this was conjectured by K.F. Roth and R. Rado, and he did not refer to matroids. Horn also acknowledged the help of Rado.

Surprisingly, the same theorem was also published by Rado [1962a] (in the same journal). The proof method (including notation) is the same as that of Horn, but no reference to Horn's paper is given. Rado wondered if the theorem can be generalized to matroids:

> It can be seen that some steps of the argument can be adapted to the more general situation of abstract independence functions but there does not appear to be an obvious way of making the whole argument apply to the more general case.

Rado [1962b] presented the vector theorem at the International Congress of Mathematicians in Stockholm in 1962, where he mentioned again that its proof has not yet been extended to 'abstract independence relations' (matroids). He wondered if the property in fact would *characterize* linear matroids.

Finally, two years later, at the Conference on General Algebra in Warsaw, 7–11 September 1964, Rado announced the base covering theorem. Simultaneously, there was the Seminar on Matroids at the National Bureau of Standards in Washington, D.C., 31 August–11 September 1964, where Edmonds [1965c] presented the base covering theorem.

In the paper based on his lecture in Warsaw, Rado [1966] did not give a proof of the matroid base cover theorem, but just said that the argument of Horn [1955] can be adapted so as to yield the more general version (as we did in Section 42.1b).

The matroid base covering theorem generalizes also the min-max relation of Nash-Williams [1964] for the minimum number of forests needed to cover the edges of a graph. (As each graphic matroid is linear, this follows also from the result of Horn [1955] described above.)

The basic unifying result (Theorem 42.1) on matroid union was given in Nash-Williams [1967], which has as special case the matroid union theorem given by Edmonds [1968]. In a footnote on page 20 of Pym and Perfect [1970], it is remarked that:

> Professor Nash-Williams has written to inform us that these results were suggested by earlier unpublished work of Professor J. Edmonds on the relation between independence structures and submodular functions.

It seems in fact much easier to prove the matroid union theorem in general, than just its special case for graphic matroids (for instance, the covering forests theorem). It also generalizes theorems of Higgins [1959] on disjoint transversals (Theorem 22.11), and of Tutte [1961a] and Nash-Williams [1961b] on disjoint spanning trees in a graph (Corollary 51.1a). (These papers mention no possible generalization to matroids.)

Welsh [1976] mentioned on these results:

> They illustrate perfectly the principle that mathematical generalization often lays bare the important bits of information about the problem at hand.

Chapter 43

Matroid matching

We saw two generalizations of Kőnig's matching theorem for bipartite graphs: the Tutte-Berge formula on matchings in arbitrary graphs and the matroid intersection theorem. This raises the demand for a common generalization of these last two theorems. A solution to the following *matroid matching problem*, posed by Lawler [1971b,1976b], could yield such a generalization: given an undirected graph $G = (S, E)$ and a matroid $M = (S, \mathcal{I})$, what is the maximum number of disjoint edges of G whose union is independent in M?

By taking M trivial, the matroid matching problem reduces to the matching problem, and by taking G regular of degree one, and M to be the disjoint sum of two matroids defined on the two colour classes of the bipartite graph G, we obtain the matroid intersection problem.

However, the general matroid matching problem has been shown to be NP-complete in the regular NP-framework, and unsolvable in polynomial time in an oracle framework.

On the other hand, Lovász [1980b] gave a strongly polynomial-time algorithm in case the matroid M is linear. Moreover, Lovász [1980a] gave a min-max relation, which was extended by Dress and Lovász [1987] to algebraic matroids.

No extension to the weighted case has been discovered, even not for the linear case: no polyhedral characterization or polynomial-time algorithm for finding a maximum-weight matroid matching has been found.

43.1. Infinite matroids

In this chapter, we need an extension of the notion of matroids to infinite matroids. An *infinite matroid* is defined as a pair $M = (S, \mathcal{I})$, where S is an infinite set and $\mathcal{I}$ is a nonempty collection of subsets of S satisfying:

(43.1)
- (i) if $I \in \mathcal{I}$ and $J \subseteq I$, then $J \in \mathcal{I}$,
- (ii) if $I \subseteq S$ and each finite subset of I belongs to $\mathcal{I}$, then I belongs to $\mathcal{I}$;
- (iii) if I, J are finite sets in $\mathcal{I}$ and $|I| < |J|$, then $I \cup \{j\} \in \mathcal{I}$ for some $j \in J \setminus I$.

Standard matroid terminology transfers to infinite matroids. The sets in $\mathcal{I}$ are called *independent* and those subsets of S not in $\mathcal{I}$ *dependent*. An inclusionwise minimal dependent set is a *circuit*. By (43.1)(ii), each *circuit* of M is finite. We will restrict ourselves to infinite matroids of *finite rank*. That is, there is a finite upper bound on the size of the sets in $\mathcal{I}$.

Examples of infinite matroids are linear spaces, where $\mathcal{I}$ is the collection of linearly independent subsets, and field extensions L of a field K, where $\mathcal{I}$ is the collection of subsets of L that are algebraically independent over K. In fact, these are the only two classes of infinite matroids that we will consider.

We call a matroid $M = (S, \mathcal{I})$ with S finite also a *finite matroid*.

43.2. Matroid matchings

Let $(S, \mathcal{I})$ be a (finite or infinite) matroid, with rank function r and span function span. Let E be a finite collection of unordered pairs from S, such that each pair is an independent set of $(S, \mathcal{I})$. For $F \subseteq E$ define

(43.2) $$\text{span}(F) := \text{span}(\bigcup F)$$

(where $\bigcup F$ denotes the union of the pairs in F), and

(43.3) $$r(F) := r(\text{span}(F)).$$

Then for $X, Y \subseteq E$ one has

(43.4) $$r(X) + r(Y) \geq r(X \cap Y) + r(X \cup Y),$$

since

(43.5) $$\begin{aligned} r(X) + r(Y) &= r(\text{span}(X)) + r(\text{span}(Y)) \\ &\geq r(\text{span}(X) \cap \text{span}(Y)) + r(\text{span}(X) \cup \text{span}(Y)) \\ &\geq r(\text{span}(X \cap Y)) + r(\text{span}(X \cup Y)) = r(X \cap Y) + r(X \cup Y). \end{aligned}$$

Call a subset M of E a *matroid matching*, or just a *matching*, if

(43.6) $$r(M) = 2|M|.$$

So M is a matroid matching if and only if M consists of disjoint pairs and the union of the pairs in M belongs to $\mathcal{I}$. Hence each subset of a matching is a matching again. The maximum size of a matching in E is denoted by $\nu(E)$, or just by ν. A matching of size $\nu(E)$ is called a *base* of E. (We should be aware of the difference between a matching in a graph and a matroid matching, and between a base of a matroid and a base of a collection of pairs in a matroid. Below we will see moreover the notion of a circuit in a set of pairs in a matroid. We will be careful to avoid confusion.[32])

Consider the function s defined on subsets F of E by

(43.7) $$s(F) := 2|F| - r(F).$$

[32] As we denote a matching by M, we denote a matroid, for the time being, just by $(S, \mathcal{I})$.

So a subset M of E is a matching if and only if $s(M) = 0$.

Then for all collections X and Y:

(43.8) (i) $s(X) \leq s(Y)$ if $X \subseteq Y$,
(ii) $s(X) + s(Y) \leq s(X \cap Y) + s(X \cup Y)$.

Here (i) follows from

(43.9) $r(Y) \leq r(X) + r(Y \setminus X) \leq r(X) + 2|Y| - 2|X|$.

(43.8)(ii) follows from (43.4).

(43.8) implies:

(43.10) each $F \subseteq E$ contains a unique inclusionwise minimal subset X with $s(X) = s(F)$.

For let F contain subsets X and Y with $s(X) = s(Y) = s(F)$. Then by (43.8)(i), $s(X \cap Y) \leq s(F)$ and $s(X \cup Y) = s(F)$, and by (43.8)(ii), $s(X \cap Y) \geq s(X) + s(Y) - s(X \cup Y) = s(F)$. So $s(X \cap Y) = s(F)$.

43.3. Circuits

A subset C of E is called a *circuit* if it is an inclusionwise minimal set satisfying $r(C) = 2|C| - 1$. By (43.10):

(43.11) each $F \subseteq E$ with $r(F) = 2|F| - 1$ contains a unique circuit.

It implies that for each matching M and each $e \in E$ with $r(M + e) = r(M) + 1$, there is a unique circuit contained in $M + e$. This circuit is denoted by $C(M, e)$, and is called a *fundamental circuit* (*of* M). (Here and below, $M + e := M \cup \{e\}$ and $M - e := M \setminus \{e\}$.)

Such circuits have a useful exchange property:

(43.12) for each $f \in C(M, e)$, $M + e - f$ is a matching again.

Indeed, if $M + e - f$ is not a matching, then $s(M + e - f) \geq 1$. In fact, $s(M + e - f) = 1$, since $s(M + e - f) \leq s(M + e) = 1$. So $M + e - f$ contains a circuit C. As $f \notin C$, we know $C \neq C(M, e)$, contradicting (43.11).

43.4. A special class of matroids

The min-max equality for matroid matching to be proved, holds for (finite or infinite) matroids $(S, \mathcal{I})$ satisfying the following condition:

(43.13) for each pair of circuits C_1, C_2 of $(S, \mathcal{I})$ with $C_1 \cap C_2 \neq \emptyset$ and $r(C_1 \cup C_2) = |C_1 \cup C_2| - 2$, the intersection of $\text{span}(C)$ taken over all circuits $C \subseteq C_1 \cup C_2$ has positive rank.

Examples of such matroids will be seen in Section 43.6.

In (43.13), 'circuits' are meant in the original meaning: as subsets of S. But the property transfers to subsets of E, as follows:

Lemma 43.1α. *Let $(S, \mathcal{I})$ be a matroid satisfying (43.13) and let E be a collection of pairs from S. Then for each pair of circuits $C_1, C_2 \subseteq E$ with $C_1 \cap C_2 \neq \emptyset$ and $s(C_1 \cup C_2) = 2$, the intersection of $\text{span}(C)$ taken over all circuits $C \subseteq C_1 \cup C_2$ has positive rank.*

Proof. Let $F := C_1 \cup C_2$. By assumption, $s(F) = 2$. Each proper subcollection F' of F satisfies $s(F') \leq 1$, since if $e \in C_i$, then $s(F - e) \leq s(F) + s(C_i - e) - s(C_i) = 2 + 0 - 1 = 1$.

Let $C_1, \ldots, C_k$ be the circuits contained in F. We can assume that $k \geq 3$ (otherwise the lemma trivially holds, since $C_1 \cap C_2 \neq \emptyset$ by assumption).

Then

(43.14) $\quad C_i \cup C_j = F$ for all distinct $i, j = 1, \ldots, k$,

since for any $e \in F \setminus (C_i \cup C_j)$ we would have that $s(F - e) = 1$ and that $F - e$ contains two distinct circuits, which contradicts (43.11).

An equivalent way of stating (43.14) is:

(43.15) $\quad F \setminus C_1, \ldots, F \setminus C_k$ are pairwise disjoint.

Now first assume that there exist distinct $e, f \in F$ with $e \cap f \neq \emptyset$. Then $|e \cup f| = 3$, so $\{e, f\}$ is a circuit, and therefore by (43.15), each C_i intersects $\{e, f\}$ (as $k \geq 3$). So each $\text{span}(C_i)$ contains $e \cap f$, and therefore the intersection of the $\text{span}(C_i)$ is nonempty, as required.

So we can assume that the pairs in F are disjoint. Consider any i. Then $\bigcup C_i$ is a subset of S, containing a unique circuit C'_i (as subset of S). This follows from:

(43.16) $\quad r(\bigcup C_i) = \left|\bigcup C_i\right| - 1$

(as C_i is a circuit in E), because (43.16) implies that $\bigcup C_i$ contains an independent set of size $\left|\bigcup C_i\right| - 1$.

Then

(43.17) $\quad C'_i \neq C'_j$ if $i \neq j$.

Indeed, C'_i intersects each pair in C_i, since for each $e \in C_i$ the union of the $f \in C_i - e$ has rank $2|C_i - e|$, hence is independent. As the pairs in F are disjoint, this shows (43.17).

Moreover, if $i \neq j$ and $h \in \{1, \ldots, k\}$, then

(43.18) $\quad C'_h \subseteq C'_i \cup C'_j$.

Otherwise, choose $x \in C'_i$, $y \in C'_j \setminus C'_i$, and $z \in C'_h \setminus (C'_i \cup C'_j)$. So $x, y, z \in \text{span}((C'_i \cup C'_j \cup C'_h) \setminus \{x, y, z\})$. Hence $r(C'_i \cup C'_j \cup C'_h) \leq |C'_i \cup C'_j \cup C'_h| - 3$, and so

(43.19) $r(F) \leq r(C'_i \cup C'_j \cup C'_h) + \left|\bigcup F\right| - |C'_i \cup C'_j \cup C'_h| \leq \left|\bigcup F\right| - 3,$

a contradiction, since $s(F) = 2$.

This proves (43.18), which implies that $C'_1 \cap C'_2 \neq \emptyset$ (since $C'_1 \subseteq C'_2 \cup C'_3$ and $C'_1 \not\subseteq C'_3$). Then by (43.13), the intersection of $\text{span}(C'_i)$ over all i has positive rank. Hence the intersection of $\text{span}(C_i)$ over all i has positive rank.

∎

For any collection E of pairs from S, let H_E be the hypergraph with vertex set E and edges all fundamental circuits. The following theorem will be used in deriving a general min-max relation.

Theorem 43.1. *Let* $(S, \mathcal{I})$ *be a matroid satisfying* (43.13) *and let* E *be a collection of pairs from* S *such that the intersection of* $\text{span}(B)$ *over all bases* B *of* E *has rank* 0. *Then*

(43.20) $|B \cap F| = \lfloor \frac{1}{2} r(F) \rfloor$

for each base B *and each component* F *of* H_E.

Proof. I. Call two fundamental circuits C, D *far* if there exist a base B and $e, g \in E$ with $r(B + e + g) = 2\nu + 2$ and with $C = C(B, e)$ and $D = C(B, g)$. We first show:

(43.21) far fundamental circuits are disjoint.

Suppose to the contrary that there exist a base B and $e, g \in E$ with $r(B + e + g) = 2\nu + 2$ and $C(B, e) \cap C(B, g) \neq \emptyset$. Let $D := C(B, e) \cup C(B, g)$. Then

(43.22) $s(D) \geq s(C(B, e)) + s(C(B, g)) - s(C(B, e) \cap C(B, g)) = 2$

and

(43.23) $s(D) \leq s(B + e + g) = 2.$

So $s(D) = 2$. If C is any circuit contained in $B + e + g$, then $C \subseteq D$, since otherwise $s(C \cap D) = 0$, and hence

(43.24) $2 = 0 + s(B + e + g) \geq s(C \cap D) + s(C \cup D) \geq s(C) + s(D) = 3,$

a contradiction.

By Lemma 43.1α, there is a nonloop p that is contained in $\text{span}(C)$ for each circuit $C \subseteq D$. By assumption, there is a base B' with $p \notin \text{span}(B')$. Choose B' with $|B' \cap (B+e+g)|$ maximal. Then $r(B'+p) = 2\nu+1 < r(B+e+g)$, and hence $f \not\subseteq \text{span}(B'+p)$ for some $f \in B+e+g$. Then $p \notin \text{span}(B'+f)$ (since $r(B' + f) \leq 2\nu + 1$), and therefore $p \notin \text{span}(C(B', f))$. So $C(B', f)$ is not one of the circuits contained in $B+e+g$. Choose $h \in C(B', f) \setminus (B+e+g)$. Hence, resetting B' to $B'-h+f$ would give a larger intersection with $B+e+g$, a contradiction. This shows (43.21).

II. We next show the theorem assuming that H_E is connected. Suppose to the contrary that $r(E) \geq 2\nu(E) + 2$. Then far fundamental circuits exist, since for any base B, there exist $e, g \in E$ with $r(B+e+g) = 2\nu+2$, since $r(E) \geq r(B)+2$. Then (43.21) implies, as H_E is connected, that there exist fundamental circuits C, C', D with C and D far, $C \cap C' \neq \emptyset$, and C' and D not far.

Choose $e \in C \cap C'$ and $f \in D$. As C and D are far fundamental circuits, there is a base B with $r(B+e+f) = 2\nu+2$ and $C = C(B,e)$, $D = C(B,f)$. Also, as C' is a fundamental circuit, there is a base B' with $r(B'+e) = 2\nu+1$ and $C' = C(B',e)$. Choose such a B' with $|B' \cap (B+f)|$ maximal.

As $r(B+e+f) > r(B'+e)$, there exists a $g \in B+f$ with $r(B'+e+g) = 2\nu+2$. As C' and D are not far, $C(B',g) \neq D = C(B,f)$. So $C(B',g) \not\subseteq B+f$, and hence there exists an $h \in C(B',g) \setminus (B+f)$. Set $B'' := B'-h+g$. Then $r(B''+h+e) = r(B'+g+e) = 2\nu+2$, and hence $r(B''+e) = 2\nu+1$. As, by (43.21), $C(B',g)$ and $C(B',e)$ are disjoint, we know $h \notin C(B',e)$, so $C(B',e) \subseteq B''+e$, and hence $C(B'',e) = C(B',e) = C'$. As $|B'' \cap (B+f)| > |B' \cap (B+f)|$ this contradicts the maximality of $|B' \cap (B+f)|$.

III. We finally prove the theorem in general. Let F be a component of H_E. Suppose that there is a base B of E with $|B \cap F| < \lfloor \frac{1}{2} r(F) \rfloor$. Then

(43.25) there is a base B of E and a base M of F with $|M| > |B \cap F|$.

Otherwise, for each base B of E, $B \cap F$ is a base of F. Then H_F consists of one component (as each fundamental circuit of E contained in F is a fundamental circuit of F). Hence, by part II of this proof, $|B \cap F| = \nu(F) = \lfloor \frac{1}{2} r(F) \rfloor$, contradicting our assumption.

So (43.25) holds. Choose B and M as in (43.25) with $|M \cap B|$ maximal. Then

(43.26) $\operatorname{span}(M) \subseteq \operatorname{span}(B)$,

since otherwise there is an $e \in M$ with $e \not\subseteq \operatorname{span}(B)$, and we can choose $f \in C(B,e) \setminus M$ and replace B by $B-f+e$, thereby increasing $|M \cap B|$, contradicting the maximality of $|M \cap B|$.

Moreover,

(43.27) for each $e \in F$ with $e \not\subseteq \operatorname{span}(B)$, we have $C(M,e) = C(B,e)$.

Otherwise, choose $f \in C(B,e) \setminus (M+e)$ and $g \in C(M,e) \setminus (B+e)$. Replacing B and M by $B-f+e$ and $M-g+e$ respectively, increases $|M \cap B|$, a contradiction.

As $M \setminus B \neq \emptyset$, there is an $h \in M \setminus B$. Then there is a base B' of E with $h \not\subseteq \operatorname{span}(B')$ (as by the condition in the theorem, there is no nonloop that is contained in the span of each base). We assume that we have chosen M, B, and B' with $|B \cap B'|$ maximal (under the primary condition that $|M \cap B|$ is maximum).

Since $h \not\subseteq \operatorname{span}(B')$, we know by (43.26) that $\operatorname{span}(B) \neq \operatorname{span}(B')$. Hence there exists an $e \in B'$ with $e \not\subseteq \operatorname{span}(B)$.

If $e \notin F$, then $C(B,e)$ is disjoint from F (as F is a component of H_E). Choose $f \in C(B,e) \setminus B'$. Then replacing B by $B - f + e$ maintains M, $B \cap F$, and $M \cap B$, but increases $|B \cap B'|$, contradicting our assumption.

So $e \in F$. By (43.27), $C(B,e) = C(M,e)$. Choose $f \in C(B,e) \setminus B'$. Then replacing M and B by $M - f + e$ and $B - f + e$ respectively, maintains $|M|$, $|B \cap F|$, and $|M \cap B|$, but increases $|B \cap B'|$, contradicting our assumption. ∎

43.5. A min-max formula for maximum-size matroid matching

We can now derive a min-max formula for the maximum size of a matching in matroids satisfying (43.13) in an hereditary way, due to Lovász [1980a]:

Theorem 43.2 (matroid matching theorem). *Let $M = (S, \mathcal{I})$ be a (finite or infinite) matroid (with rank function r) such that each contraction of M satisfies* (43.13). *Let E be a finite set of pairs from S. Then the maximum size $\nu(E)$ of a matching in E satisfies*

$$\nu(E) = \min(r(F) + \sum_{i=1}^{k} \lfloor \tfrac{1}{2}(r(F_i) - r(F)) \rfloor), \tag{43.28}$$

where $F, F_1, \ldots, F_k$ are flats such that $F \subseteq F_i$ for $i = 1, \ldots, k$, and such that each $e \in E$ is contained in some F_i.

Proof. We first show that $\leq$ holds in (43.28). Let B be a base of E, and partition B into $B_1, \ldots, B_k$ such that $\text{span}(B_i) \subseteq F_i$ for $i = 1, \ldots, k$. Define $F'_i := \text{span}(B_i)$.

By induction on l we show that for each $l = 0, \ldots, k$:

$$r(F \cup F'_1 \cup \cdots \cup F'_l) \leq r(F) + \sum_{i=1}^{l} (|B_i| + \lfloor \tfrac{1}{2}(r(F \cup F'_i) - r(F)) \rfloor). \tag{43.29}$$

For $l = 0$ this is trivial. For $l \geq 1$ we have (by induction and submodularity):

$$\begin{aligned} r(F \cup F'_1 \cup \cdots \cup F'_l) &\leq r(F \cup F'_1 \cup \cdots \cup F'_{l-1}) + r(F \cup F'_l) - r(F) \\ &\leq r(F \cup F'_l) + \sum_{i=1}^{l-1} (|B_i| + \lfloor \tfrac{1}{2}(r(F \cup F'_i) - r(F)) \rfloor) \\ &\leq r(F) + \sum_{i=1}^{l} (|B_i| + \lfloor \tfrac{1}{2}(r(F \cup F'_i) - r(F)) \rfloor), \end{aligned} \tag{43.30}$$

since

$$r(F \cup F'_l) \leq r(F) + |B_l| + \tfrac{1}{2}(r(F \cup F'_l) - r(F)), \tag{43.31}$$

as $|B_l| = \frac{1}{2}r(F'_l)$. This shows (43.29), which for $l = k$ implies that $\nu(E)$ is at most (43.28), since

$$(43.32)\qquad \begin{aligned} 2\nu(E) &\le r(F \cup F'_1 \cup \cdots \cup F'_k) \\ &\le r(F) + \sum_{i=1}^{k} \big(|B_i| + \lfloor \tfrac{1}{2}(r(F \cup F'_i) - r(F))\rfloor\big) \\ &= \nu(E) + r(F) + \sum_{i=1}^{l} \lfloor \tfrac{1}{2}(r(F \cup F_i) - r(F))\rfloor. \end{aligned}$$

Equality is shown by induction on $r(M)$. First assume that there is a nonloop p that is contained in $\mathrm{span}(B)$ for each base B of E. Let M' be the matroid M/p obtained by contracting p. Let E' be the set of pairs $\{s,t\}$ in E such that $s, t \neq p$ and such that s and t are not parallel in M'. Let ν' be the maximum size of a base $B' \subseteq E'$ with respect to M'.

Then $\nu' < \nu(E)$. For suppose that $\nu' \geq \nu(E)$. Let B' be a base of E' with respect to M'. As $|B'| \geq \nu(E)$, B' is also a base of E with respect to M. As $r_{M'}(B') = 2|B'| = r_M(B')$, we have $p \notin \mathrm{span}_M(B)$. This contradicts our assumption.

So $\nu' < \nu(E)$. By induction, M' has flats $F', F'_1, \ldots, F'_{k'}$ with $F' \subseteq F'_i$ for $i = 1, \ldots, k'$, such that each $e \in E'$ is contained in some F'_i and such that

$$(43.33)\qquad \nu' = r_{M'}(F') + \sum_{i=1}^{k'} \lfloor \tfrac{1}{2}(r_{M'}(F'_i) - r_{M'}(F'))\rfloor.$$

Define $F := \mathrm{span}_M(F' + p)$ and $F_i := \mathrm{span}_M(F'_i + p)$ for $i = 1, \ldots, k'$. Moreover, for each $e \in E$ not occurring in E', introduce a new F_i with $F_i := \mathrm{span}_M(F + e)$. As $p \in F$, we have $r_M(F_i) \leq r_M(F) + 1$ for each of these F_i.

This gives $F, F_1, \ldots, F_k$ such that $F \subseteq F_i$ for $i = 1, \ldots, k$, such that each $e \in E$ is contained in some F_i and such that

$$(43.34)\qquad \begin{aligned} \nu(E) &\geq \nu' + 1 = r_{M'}(F') + 1 + \sum_{i=1}^{k'} \lfloor \tfrac{1}{2}(r_{M'}(F'_i) - r_{M'}(F'))\rfloor \\ &= r(F) + \sum_{i=1}^{k} \lfloor \tfrac{1}{2}(r(F_i) - r(F))\rfloor. \end{aligned}$$

So we can assume that there is no nonloop p contained in $\mathrm{span}(B)$ for all bases B of E. Let $E_1, \ldots, E_k$ be the components of H_E and let $F_i := \mathrm{span}(E_i)$ for $i = 1, \ldots, k$. Let B be a base of E. Then by (43.20),

$$(43.35)\qquad \nu(E) = |B| = \sum_{i=1}^{k} |B \cap E_i| = \sum_{i=1}^{k} \lfloor \tfrac{1}{2} r(F_i)\rfloor.$$

So taking $F := \emptyset$ gives (43.28). ∎

43.6. Applications of the matroid matching theorem

We now consider specific classes of matroids satisfying (43.13), such that we know that the min-max equality holds. First, the linear matroids (Lovász [1980b]):

Corollary 43.2a. *If E is a finite set of pairs from a linear space S, then* (43.28) *holds, where flats are linear subspaces of S.*

Proof. Let $\mathcal{I}$ be the collection of sets of linearly independent vectors in S. We must show that each contraction of the infinite matroid $M = (S, \mathcal{I})$ satisfies (43.13). It suffices to show that M satisfies (43.13), since each contraction of M is again coming from a linear space, up to loops and parallel elements.

Let C_1 and C_2 be intersecting circuits in M with $r(C_1 \cup C_2) = |C_1 \cup C_2| - 2$. As C_1 is a circuit, there is a nonzero vector p in $\mathrm{span}(C_1 \setminus C_2) \cap \mathrm{span}(C_1 \cap C_2)$, since $r(C_1 \setminus C_2) + r(C_1 \cap C_2) > r(C_1)$. Consider any circuit C contained in $C_1 \cup C_2$.

Suppose $p \notin \mathrm{span}(C)$. As $p \in \mathrm{span}(C_1 \setminus C_2) \cap \mathrm{span}(C_1 \cap C_2)$, C misses an element $s \in C_1 \setminus C_2$ and an element $t \in C_1 \cap C_2$. Now $t \in \mathrm{span}(C_2 - t)$ and $s \in \mathrm{span}(C_1 - s)$. Hence $(C_1 \cup C_2) - s - t$ spans $C_1 \cup C_2$, and hence, as $C_1 \cup C_2$ has rank $|C_1 \cup C_2| - 2$, we have that $(C_1 \cup C_2) - s - t$ is independent. This contradicts the fact that C is contained in $(C_1 \cup C_2) - s - t$. ∎

Dress and Lovász [1987] proved that a similar result holds for algebraic dependence in field extensions (where $\mathrm{tr}_K(E)$ denotes the transcendence degree of $\bigcup E$ over K):

Corollary 43.2b. *Let E be a finite set of pairs from a field extension L of a field K. Then the maximum number of disjoint pairs from E such that the union is algebraically independent over K is equal to the minimum value of*

$$\text{(43.36)} \qquad \mathrm{tr}_K(F) + \sum_{i=1}^{k} \lfloor \tfrac{1}{2} \mathrm{tr}_F(E_i) \rfloor,$$

where F ranges over all field extensions of K in L and where $E_1, \ldots, E_k$ ranges over all partitions of E.

Proof. Let $M = (L, \mathcal{I})$ be the infinite matroid with $\mathcal{I}$ consisting of all subsets of L that are algebraically independent over K.

Similarly as for the previous corollary, it suffices to show that for any two intersecting circuits C_1 and C_2 of M with $r(C_1 \cup C_2) = |C_1 \cup C_2| - 2$ there is an $\alpha \in L \setminus \mathrm{span}(K)$ such that α belongs to $\mathrm{span}(C)$ for each circuit C contained in $C_1 \cup C_2$.

Let $I := C_1 \setminus C_2$. Then

(43.37) $\qquad I$ is a circuit in M/C_2.

To see this, trivially I is dependent in M/C_2. Consider any circuit $C \subseteq C_1 \cup C_2$ intersecting I. We must show that $I \subseteq C$. Suppose that there is an $s \in I \setminus C$. As C intersects I, C misses at least one element of C_2, say t. So $C \subseteq (C_1 \cup C_2) - s - t$. Now $(C_1 \cup C_2) - s - t$ spans $C_1 \cup C_2$ (since $t \in \text{span}(C - t)$ and $s \in \text{span}(C_1 \cup C_2 - s)$). This implies that $(C_1 \cup C_2) - s - t$ is independent (as $r(C_1 \cup C_2) = |C_1 \cup C_2| - 2$), contradicting the fact that it contains a circuit. This proves (43.37).

Let $I = \{s_1, \ldots, s_n\}$. Since I is a circuit in M/C_2, there exists an irreducible polynomial p in $\text{span}(C_2)[x_1, \ldots, x_n]$ with $p(s_1, \ldots, s_n) = 0$. We can choose p such that at least one coefficient of p equals 1. Note that p has at least one coefficient, α say, that is not in $\text{span}(K)$, since I is independent over K. It therefore is enough to show that all coefficients of p belong to $\text{span}(C)$ for each circuit C contained in $C_1 \cup C_2$, since then α belongs to each $\text{span}(C)$.

Choose a circuit $C \neq C_2$ with $C \subseteq C_1 \cup C_2$. As I is a circuit in M/C_2, we have $C \setminus C_2 = I$. So I is a circuit in $M/(C \cap C_2)$. Hence there exists an irreducible polynomial q in $\text{span}(C \cap C_2)[x_1, \ldots, x_n]$ with $q(s_1, \ldots, s_n) = 0$. As $\text{span}(C \cap C_2)$ is algebraically closed in $\text{span}(C_2)$, q is also irreducible in $\text{span}(C_2)[x_1, \ldots, x_n]$[33]. Then p and q are also irreducible in $\text{span}(C_2)(x_1, \ldots, x_{n-1})[x_n]$ (cf., for instance, Section IV:6 of Jacobson [1951]). Therefore, p is a multiple of q in $\text{span}(C_2)(x_1, \ldots, x_{n-1})$; that is, there are nonzero $r, s \in \text{span}(C_2)[x_1, \ldots, x_{n-1}]$ with $rp = sq$. Hence by the unique factorization theorem (cf., for instance, Section IV:6 of Jacobson [1951]), $p = \lambda q$ for some $\lambda \in \text{span}(C_2)$. As some coefficient of p equals 1, $\lambda \in \text{span}(C \cap C_2)$. Hence $p \in \text{span}(C \cap C_2)[x_1, \ldots, x_n]$. ∎

(The property of algebraic matroids shown in this proof generalizes a property shown by Ingleton and Main [1975].)

We also formulate the special case of graphic matroids:

Corollary 43.2c. *Let $G = (V, E)$ be a graph and let $\mathcal{P}$ be a partition of E into pairs. Then the maximum size of a forest $F \subseteq E$ that is the union of classes of $\mathcal{P}$ is equal to the minimum value of*

$$(43.38) \qquad 2|V| - 2|\mathcal{Q}| + 2\sum_{i=1}^{k} \lfloor \tfrac{1}{2}\delta_{\mathcal{Q}}(E_i) \rfloor,$$

[33] This can be seen as follows. Let L be a field extension of field K, such that K is algebraically closed in L. Then if p is an irreducible polynomial in $K[x_1, \ldots, x_n]$, then p is irreducible also in $L[x_1, \ldots, x_n]$. For suppose to the contrary that $p = p_1 p_2$ for nonconstant polynomials p_1, p_2 in $L[x_1, \ldots, x_n]$. We can assume that p_1 has at least one coefficient in K. Hence, as p is irreducible in $K[x_1, \ldots, x_n]$, p_1 has at least one coefficient not in K. Choose a large enough natural number k such that substituting x_i by x^{k^i} for $i = 1, \ldots, n$, transforms p_1 to a polynomial $\tilde{p}_1$ in $L[x] \setminus K[x]$. Let $\tilde{p} \in K[x]$ be obtained similarly from p. Now the algebraic closure of K contains all roots of $\tilde{p}$, hence all roots of $\tilde{p}_1$, and hence all coefficients of $\tilde{p}_1$. As each element in $L \setminus K$ is transcendental over K, we have a contradiction.

where $\mathcal{Q}$ ranges over partitions of V into nonempty classes and where $E_1, \ldots, E_k$ ranges over partitions of E such that each E_i is a union of pairs in $\mathcal{P}$. In (43.38), $\delta_{\mathcal{Q}}(E_i)$ denotes the size of a largest forest in the graph obtained from (V, E_i) by contracting each class in $\mathcal{Q}$ to one vertex.

Proof. We apply Theorem 43.2 to the cycle matroid M of the graph H obtained from the complete graph on V by adding a parallel edge for each edge in E. Then (43.13) is satisfied for each contraction of M.

Now for each flat F of M there is a partition $\mathcal{Q}$ of V such that F is the set of edges of H contained in a class of $\mathcal{Q}$. The rank $r(F)$ of F (in M) is equal to $|V| - |\mathcal{Q}|$. For any $E' \subseteq E$, the smallest flat F' containing $F \cup E'$ has rank $r(F') = \delta_{\mathcal{Q}}(E') + r(F)$. Hence the corollary follows from Theorem 43.2. ∎

This corollary implies the following result on 3-uniform hypergraphs. A *hypergraph* is a pair $H = (V, \mathcal{E})$, where V is a finite set and $\mathcal{E}$ is a family of subsets of V. The hypergraph is called *k-uniform* if $|U| = k$ for each $U \in \mathcal{E}$.

A subfamily $\mathcal{F}$ of $\mathcal{E}$ is called a *forest* if there do not exist distinct $v_1, \ldots, v_t \in V$ and distinct $U_1, \ldots, U_t \in \mathcal{F}$ such that $t \geq 2$ and $v_{i-1}, v_i \in U_i$ for $i = 1, \ldots, t$, setting $v_0 := v_t$.

Corollary 43.2c implies a min-max relation for the maximum size of a forest in a given 3-uniform hypergraph (Lovász [1980a]):

Corollary 43.2d. *Let $H = (V, \mathcal{E})$ be a 3-uniform hypergraph. Then the maximum size of a forest $\mathcal{F} \subseteq \mathcal{E}$ is equal to the minimum value of*

(43.39) $$|V| - |\mathcal{Q}| + \sum_{\mathcal{S} \in \Sigma} \lfloor \tfrac{1}{2}(\phi_{\mathcal{Q}}(\mathcal{S}) - 1) \rfloor,$$

where $\mathcal{Q}$ and Σ range over partitions of V and $\mathcal{E}$, respectively. Here $\phi_{\mathcal{Q}}(\mathcal{S})$ denotes the number of classes of $\mathcal{Q}$ intersected by $\bigcup \mathcal{S}$.

Proof. For each $U \in \mathcal{E}$, choose two different pairs $e_U, f_U \subseteq U$, and let $G = (V, E)$ be the graph with edges all e_U and f_U. Let $\mathcal{P}$ be the partition of E into the pairs e_U, f_U. Then the maximum size of a forest $\mathcal{F} \subseteq \mathcal{E}$ is equal to half of the maximum size of a forest in E that is the union of pairs in $\mathcal{P}$. So to see that Corollary 43.2c implies the present corollary, it suffices to show that minimum (43.39) is equal to half of minimum (43.38).

First, let $\mathcal{Q}$ and Σ attain minimum (43.39). The partition Σ of $\mathcal{E}$ induces a partition of E into classes $\{e_U, f_U \mid U \in \mathcal{S}\}$ for $\mathcal{S} \in \Sigma$. One easily checks that for each $\mathcal{S} \in \Sigma$:

(43.40) $$\delta_{\mathcal{Q}}(\{e_U, f_U \mid U \in \mathcal{S}\}) \leq \phi_{\mathcal{Q}}(\mathcal{S}) - 1,$$

which implies that the minimum (43.39) is not less than half of minimum (43.38).

Second, to see the reverse inequality, let $\mathcal{Q}$, $E_1, \ldots, E_k$ attain minimum (43.38). Consider any $i = 1, \ldots, k$. Let $\mathcal{Q}'$ be the set of those classes in $\mathcal{Q}$

intersected by E_i and let t be the number of components of the hypergraph $(V, \mathcal{Q}' \cup E_i)$. Then $\delta_{\mathcal{Q}}(E_i) = |\mathcal{Q}'| - t$. The components partition E_i into $E_{i,1}, \ldots, E_{i,t}$. Then

$$\delta_{\mathcal{Q}}(E_i) = |\mathcal{Q}'| - t = \sum_{j=1}^{t} (\phi_{\mathcal{Q}}(E_{i,j}) - 1). \tag{43.41}$$

So letting Σ to be the partition of $\mathcal{E}$ into classes $\mathcal{S}_{i,j} := \{U \mid e_U, f_U \in E_{i,j}\}$ (for all i, j), we have that minimum (43.38) is not less than twice minimum (43.39). ∎

(Szigeti [1998a] gave a direct proof of this theorem for the case where the hypergraph consists of all triangles of a given graph.)

Other applications of matroid matching are a derivation of Mader's theorem on maximum packings of T-paths (cf. Chapter 73), to rigidity (see Lovász [1980a]), and to matching forests (an easy application, see Section 59.6b).

43.7. A Gallai theorem for matroid matching and covering

We prove a Gallai-type theorem that relates the maximum size of a matroid matching to the minimum number of pairs spanning the matroid.

Let E be a collection of pairs of elements from a matroid $(S, \mathcal{I})$ such that each pair is an independent set and such that $\mathrm{span}(E) = S$. Call $F \subseteq E$ a *matroid cover* if $\mathrm{span}(F) = S$. Let $\rho(E)$ be the minimum size of a matroid cover. The following relation between $\nu(E)$ and $\rho(E)$ was observed by Lovász and extends Gallai's theorem (Theorem 19.1):

Theorem 43.3. *Let $(S, \mathcal{I})$ be a matroid, with rank function r, and let E be a collection of pairs from S spanning S. Then $\nu(E) + \rho(E) = r(S)$.*

Proof. To see $\leq$, let M be matching of size $\nu(E)$. Then by adding at most $r(S) - r(M)$ pairs from E to M we obtain a matroid cover F. So $\rho(E) \leq |F| \leq |M| + (r(S) - r(M)) = \nu(E) + r(S) - 2\nu(E) = r(S) - \nu(E)$.

To see $\geq$, let F be a matroid cover of size $\rho(E)$. Let $M := F$. As long as M contains an element e with $r(M - e) \geq r(M) - 1$, delete e from M. We end up with a matching M. For suppose not. Let M' be a maximum-size matching contained in M, and choose $e \in M \setminus M'$. Then $r(M-e) \leq r(M) - 2$ (otherwise we would delete e from M). Hence:

$$r(M' + e) \geq r(M') + r(M) - r(M - e) \geq r(M') + 2 = 2|M'| + 2. \tag{43.42}$$

So $M' + e$ is a matching, contradicting the maximality of M'.

So M is a matching. Each time we have deleted an edge from M, its rank drops by at most 1. Hence $r(M) \geq r(S) - (|F| - |M|)$. Therefore $\nu(E) \geq |M| = r(M) - |M| \geq r(S) - |F| = r(S) - \rho(E)$. ∎

This theorem implies that formula (43.28) for the maximum size of a matching yields a formula for the minimum number of lines spanning all space.

43.8. Linear matroid matching algorithm

Jensen and Korte [1982] and Lovász [1981] showed that no polynomial-time algorithm exists for the matroid matching problem in general (see Section 43.9). On the other hand, Lovász [1981] gave a strongly polynomial-time algorithm for the matroid matching problem for linear matroids (an explicit representation over a field is required). This extends, e.g., Edmonds' polynomial-time algorithm finding a maximum matching in an undirected graph (cf. Section 24.2). It does not extend Edmonds' algorithm for a maximum-size common independent set in two matroids, as this algorithm also works for nonlinear matroids.

Theorem 43.4. *Given a set E of pairs of vectors in a linear space L, a maximum-size matching can be found in strongly polynomial time.*

Proof. The algorithm is a 'brute-force' polynomial-time algorithm, based on collecting many matchings and utilizing standard linear-algebraic operations, which can be performed in strongly polynomial time. Since we deal with subsets of a vector space, we can use $X + Y := \{x + y \mid x \in X, y \in Y\}$. For each $X \subseteq L$, $\text{span}(X)$ is a subspace of L.

Throughout this proof, $\mathcal{B}$ will be a collection of matchings, all of the same size ν (say). Define:

(43.43) $K_{\mathcal{B}} := \bigcap\{\text{span}(B) \mid B \in \mathcal{B}\}$ and $H_{\mathcal{B}} :=$ the hypergraph with vertex set E and edges all fundamental circuits of all $B \in \mathcal{B}$.

We say that we *improve* $\mathcal{B}$ if we find, in strongly polynomial time, either a matching B of size $\nu + 1$, or of size ν such that $K_{\mathcal{B}} \not\subseteq \text{span}(B)$, or of size ν such that $H_{\mathcal{B}\cup\{B\}}$ has fewer components than $H_{\mathcal{B}}$. So replacing $\mathcal{B}$ by $\{B\}$ if $|B| = \nu + 1$, and by $\mathcal{B} \cup \{B\}$ if $|B| = \nu$, we can have at most $2|E|$ improvements.

I. We first show (where a component is called *nontrivial* if it has more than one element):

(43.44) We can improve $\mathcal{B}$ if we have a union F of nontrivial components of $H_{\mathcal{B}}$, a matching $M \subseteq F$, and a $B \in \mathcal{B}$ such that $r(M \cup A) > |B \cap F| + |M|$, where $A := \text{span}(B \cap F) \cap K_{\mathcal{B}}$.

Here and below, $r(X \cup Y) := r(\bigcup X \cup Y)$ for $X \subseteq E$ and $Y \subseteq S$.

To see (43.44), apply the first applicable case of the following five cases, and then iterate. If Case 1 applies, we improve $\mathcal{B}$. In any of the other cases, we reset B or M or both, add the reset B to $\mathcal{B}$, and iterate with the reset B and M. The input condition given in (43.44) is maintained, as will be shown after describing the five cases.

Case 1: There is a $B' \in \mathcal{B}$ and an $e \in E$ such that $B' + e$ is a matching, or such that $C(B', e)$ intersects both F and $E \setminus F$, or such that $K_{\mathcal{B}} \not\subseteq \text{span}(B' - f + e)$ for some $f \in C(B', e)$. Output $B' + e$, B', or $B' - f + e$ (thus we improve $\mathcal{B}$).

Note that if Case 1 does not apply, then

(43.45) $\quad f \not\subseteq K_{\mathcal{B}}$ for each $f \in F$.

Indeed, as f is in a nontrivial component of $H_{\mathcal{B}}$, f is contained in some fundamental circuit $C(B', e)$ for some $B' \in \mathcal{B}$. As Case 1 does not apply, we know $K_{\mathcal{B}} \subseteq \text{span}(B' - f + e)$. Hence, if $f \subseteq K_{\mathcal{B}}$, then $f \subseteq \text{span}(B' - f + e)$, hence $2\nu + 1 = r(B' + e) = r(B' - f + e) = 2\nu$, a contradiction.

Case 2: There is an $e \in F$ such that $M + e$ is a matching and $r((M \cup A) + e) \geq r(M \cup A) + 1$. Reset $M := M + e$.

Case 3: $\text{span}(M) \not\subseteq \text{span}(B)$. Choose $e \in M$ with $e \not\subseteq \text{span}(B)$, choose $f \in C(B, e) \setminus M$, and reset $B := B - f + e$.

Case 4: There is an $e \in F$ such that $e \not\subseteq \text{span}(B)$ and $C(B, e) \neq C(M, e)$. (Note: $e \not\subseteq \text{span}(M) + A$, since $\text{span}(M) + A \subseteq \text{span}(B)$ (as Case 3 does not apply). So, as Case 2 does not apply, $M + e$ is not a matching. Hence $C(M, E)$ is defined.)

Choose $f \in C(B, e) \setminus (M + e)$ and $g \in C(M, e) \setminus (B + e)$, and reset $B := B - f + e$ and $M := M - g + e$.

Case 5. Choose $B' \in \mathcal{B}$ with $\text{span}(M \triangle (B \cap F)) \not\subseteq \text{span}(B')$ and with $|B \cap B'|$ maximal. (This is possible, since $M \neq B \cap F$, since $r((B \cap F) \cup A) = r(B \cap F) = 2|B \cap F|$ and $r(M \cup A) > |B \cap F| + |M|$ by assumption. As $M \triangle (B \cap F) \subseteq F$, such a B' exists, by (43.45).)

Choose $e \in B'$ with $e \not\subseteq \text{span}(B)$. (This is possible since $\text{span}(M \triangle (B \cap F)) \subseteq \text{span}(B)$, so $\text{span}(B) \neq \text{span}(B')$.)

Choose $f \in C(B, e) \setminus B'$. If $e \notin F$, reset $B := B - f + e$. If $e \in F$, reset $B := B - f + e$ and $M := M - f + e$. (Note that if $e \in F$, then $C(B, e) = C(M, e)$ as Case 4 does not apply.)

Running time. The number of iterations is polynomially bounded, since in each iteration (except the last, where Case 1 applies), the vector $(|M|, |M \cap B|, |B \cap B'|)$ increases lexicographically. Here it is important to note that Case 5 does not modify the set $M \triangle (B \cap F)$, and increases the intersection of this set with B'.

We finally prove that the resettings in Cases 2-5 indeed maintain the condition given in (43.44). Let $\widetilde{B}$, $\widetilde{M}$, and $\widetilde{A}$ denote B, M, and A after resetting (taking $\widetilde{B}$ or $\widetilde{M}$ equal to B or M if they are not reset). We must show

(43.46) $\quad r(\widetilde{M} \cup \widetilde{A}) > |\widetilde{B} \cap F| + |\widetilde{M}|.$

Note that, as Case 1 does not apply, $|\widetilde{B} \cap F| = |B \cap F|$.

We first show:

(43.47) $A \subseteq \widetilde{A}$.

This is equivalent to (since $K_{\mathcal{B}}$ does not change, as Case 1 does not apply):

(43.48) $A \subseteq \text{span}(\widetilde{B} \cap F)$.

This is trivial if $\widetilde{B} \cap F = B \cap F$. So we can assume that $\widetilde{B} \cap F \neq B \cap F$. Hence $\widetilde{B} = B - f + e$ for some $e, f \in F$. Then (43.48) follows from

(43.49)
$$\begin{aligned} r((\widetilde{B} \cap F) \cup A) &\leq r((\widetilde{B} \cap F) \cup A + f) - 1 = r((B \cap F) \cup A + e) - 1 \\ &= r((B \cap F) + e) - 1 \leq r(B \cap F) = 2|B \cap F| = 2|\widetilde{B} \cap F| \\ &= r(\widetilde{B} \cap F). \end{aligned}$$

Here the first inequality holds as $f \not\subseteq \text{span}(\widetilde{B} \cap F) + A$, since $f \not\subseteq \text{span}(\widetilde{B})$ and $\text{span}(\widetilde{B} \cap F) + A \subseteq \text{span}(\widetilde{B})$. (We use that $A \subseteq K_{\mathcal{B}} \subseteq \text{span}(\widetilde{B})$, as Case 1 does not apply.) The last inequality holds as $(B \cap F) + e$ is not a matching, since it contains $C(B, e)$ (as Case 1 does not apply). This shows (43.48), and hence (43.47).

We finally show (43.46). In Case 2, we have $\widetilde{B} = B$, $\widetilde{M} = M + e$, and $\widetilde{A} = A$, and hence

(43.50)
$$\begin{aligned} r(\widetilde{M} \cup \widetilde{A}) &= r((M \cup A) + e) \geq r(M \cup A) + 1 > |B \cap F| + |M| + 1 \\ &= |\widetilde{B} \cap F| + |\widetilde{M}|, \end{aligned}$$

as required.

In Case 3, (43.47) implies (as $\widetilde{M} = M$) that $r(\widetilde{M} \cup \widetilde{A}) \geq r(M \cup A) > |B \cap F| + |M| = |\widetilde{B} \cap F| + |M|$.

In Cases 4 and 5 we have $\widetilde{M} = M - g + e$ (possible $g = f$). Then

(43.51)
$$\begin{aligned} r(\widetilde{M} \cup \widetilde{A}) &\geq r(\widetilde{M} \cup A) \geq r((\widetilde{M} \cup A) + g) - 1 = r((M \cup A) + e) - 1 \\ &\geq r(M \cup A) > |B \cap F| + |M| = |\widetilde{B} \cap F| + |\widetilde{M}|. \end{aligned}$$

The first inequality follows from (43.47). Next, $e \not\subseteq \text{span}(M \cup A)$ (as $e \not\subseteq \text{span}(B)$ and as $A \subseteq \text{span}(B)$ and $\text{span}(M) \subseteq \text{span}(B)$, since Case 3 does not apply). This gives the third inequality. To see the second inequality, suppose it does not hold. Then $\widetilde{M} + g$ is a matching, hence $M + e$ is a matching. Therefore, as Case 2 does not apply, $r(M \cup A + e) = r(M \cup A)$, contradicting the fact that $e \not\subseteq \text{span}(M \cup A)$.

II. Secondly,

(43.52) we can improve $\mathcal{B}$ if $K_{\mathcal{B}} = \{\mathbf{0}\}$, $H_{\mathcal{B}}$ is connected, and $\nu < \lfloor \frac{1}{2} r(E) \rfloor$.

(In this case, $\mathcal{B}$ can only be improved by finding a matching larger than B.)

The algorithm follows the framework of parts I and II in the proof of Theorem 43.1. Again, the algorithm iteratively applies the first applicable case. Call two circuits C_1, C_2 *far* if there exist $B \in \mathcal{B}$ and $e, g \in E$ with $r(B + e + g) = 2\nu + 2$ and $C_1 = C(B, e)$ and $C_2 = C(B, g)$.

Case 1: There exists a $B \in \mathcal{B}$ and $e \in E$ such that $B + e$ is a matching of size $\nu + 1$. Output $B + e$.

Case 2: There exist far circuits C_1 and C_2 with $C_1 \cap C_2 \neq \emptyset$. We will create a matching of size $\nu + 1$.

Let $C_1 = C(B, e)$ and $C_2 = C(B, g)$ for some $B \in \mathcal{B}$ with $e, g \in E$ and $r(B+e+g) = 2\nu+2$. Define $D := C_1 \cup C_2$. As is shown in the proof of Corollary 43.2a, there is a $p \neq \mathbf{0}$ contained in $\text{span}(C)$ for each circuit $C \subseteq D$. Since $K_{\mathcal{B}} = \{\mathbf{0}\}$, there is a $B' \in \mathcal{B}$ with $p \notin \text{span}(B')$. Now $r(B' + p) = 2\nu + 1 < r(B + e + g)$, and hence $f \not\subseteq \text{span}(B' + p)$ for some $f \in B + e + g$. Then (as $B' + f$ is not a matching, since Case 1 does not apply) $p \notin \text{span}(B' + f)$, and therefore $p \notin \text{span}(C(B', f))$. So $C(B', f)$ is not contained in $B + e + g$. Choose $h \in C(B', f) \setminus (B + e + g)$. Hence, resetting B' to $B' - h + f$ increases $|B' \cap (B + e + g)|$. So iterating this, we finally obtain a matching larger than ν.

Case 3. We show that we can create a matching of size $\nu + 1$, or make that Case 1 or 2 applies.

Far circuits exist, since for any base B, there exist $e, g \in E$ with $r(B+e+g) = 2\nu+2$, since $r(E) \geq r(B)+2$. Choose far circuits C, D that are closest[34] together in the hypergraph $H_{\mathcal{B}}$. Assuming that Case 2 does not apply, we know $C \cap D = \emptyset$. Hence there is an intermediate set C' on a shortest path from C to D. Let $C = C(B, e)$, $D = C(B, g)$, and $C' = C(B', f)$ for $B, B' \in \mathcal{B}$ and $e, f, g \in E$ with $r(B+e+g) = 2\nu+2$. We choose B' such that $|B' \cap (B+e+g)|$ is maximal. Choose $h \in B + e + g$ with $h \not\subseteq \text{span}(B' + f)$.

$C(B', h)$ and $C(B', f)$ are disjoint, since otherwise we can apply Case 2. Moreover,

(43.53) $\quad C(B', h) \not\subseteq B + e + g,$

Otherwise, $C(B', h) = C(B, e)$ or $C(B', h) = C(B, g)$. Hence C' and C or D are far, contradicting the minimality of the distance of C and D.

Hence we have (43.53). Choose $i \in C(B', h) \setminus (B + e + g)$ and add $B'' := B' - i + h$ to $\mathcal{B}$. Iterate Case 3 with B' replaced by B'' (note that $C' = C(B'', f)$). As $|B'' \cap B| > |B' \cap B|$, the number of iterations of Case 3 is at most ν.

III. Combination of the previous two algorithms implies:

(43.54) $\quad$ we can improve $\mathcal{B}$ if $K_{\mathcal{B}} = \{\mathbf{0}\}$ and $\nu < \nu(E)$.

As $\nu < \nu(E)$, there is a component F of $H_{\mathcal{B}}$ with $|B \cap F| < \nu(F) \leq \lfloor \frac{1}{2} r(F) \rfloor$ for at least one $B \in \mathcal{B}$. If there exist $B, B' \in \mathcal{B}$ with $|B \cap F| < |B' \cap F|$, set

[34] Here the distance of fundamental circuits C, D is the minimum length of a path connecting C and D. A *path* connecting C and D is a sequence $C = C_0, \ldots, C_k = D$ of fundamental circuits such that $C_{i-1} \cap C_i \neq \emptyset$ for $i = 1, \ldots, k$. Its length is k.

$M := B' \cap F$. Otherwise (that is, if $|B \cap F| = |B' \cap F|$ for all $B' \in \mathcal{B}$), apply (43.52) to $\mathcal{B}' := \{B \cap F \mid B \in \mathcal{B}\}$ and $B \cap F$ for any $B \in \mathcal{B}$, to obtain a matching $M \subseteq F$ with $|M| = |B \cap F| + 1$.

Now applying (43.44) to $\mathcal{B}$, F, B, and M improves $\mathcal{B}$. (Since $A \subseteq K_{\mathcal{B}}$, we have $A = \{\mathbf{0}\}$, and hence $r(M \cup A) = r(M) = 2|M| > |M| + |B \cap F|$.)

IV. Finally:

(43.55) We can improve $\mathcal{B}$ if $\mathcal{B} \neq \emptyset$ and $\nu < \nu(E)$.

Define F to be the union of all fundamental circuits of the $B \in \mathcal{B}$. This implies

(43.56) $\operatorname{span}(E \setminus F) \subseteq K_{\mathcal{B}}$.

If there exist $B, B' \in \mathcal{B}$ with $|B \cap F| < |B' \cap F|$, then applying (43.44) to B and $M := B' \cap F$ improves $\mathcal{B}$. So we can assume that $|B \cap F| = \beta$ for all $B \in \mathcal{B}$. Choose $B_0 \in \mathcal{B}$ with $r(\operatorname{span}(B_0 \cap F) \cap K_{\mathcal{B}})$ maximal. Define

(43.57) $A := \operatorname{span}(B_0 \cap F) \cap K_{\mathcal{B}}$, $E' := F/A$, and $\nu' := \beta - r(A)$.

For each $B \in \mathcal{B}$ there is a matching M_B in $(B \cap F)/A$ of size ν', since

(43.58)
$$\begin{aligned}&\beta - r(\operatorname{span}(B \cap F) \cap A) \geq \beta - r(\operatorname{span}(B \cap F) \cap K_{\mathcal{B}})\\ &\geq \beta - r(\operatorname{span}(B_0 \cap F) \cap K_{\mathcal{B}}) = \beta - r(A) = \nu'.\end{aligned}$$

Let $\mathcal{B}' := \{M_B \mid B \in \mathcal{B}\}$. Then $K_{\mathcal{B}'} = \{\mathbf{0}\}$, since

(43.59)
$$\begin{aligned}&\bigcap_{B \in \mathcal{B}} \operatorname{span}(B \cap F) \subseteq \operatorname{span}(B_0 \cap F) \cap \bigcap_{B \in \mathcal{B}} \operatorname{span}(B)\\ &= \operatorname{span}(B_0 \cap F) \cap K_{\mathcal{B}} = A.\end{aligned}$$

Since $\nu(E) > \nu$ we have $\nu(E') > \nu'$. Indeed, let B' be a matching in E of size $\nu + 1$. Then B'/A contains a matching of size $|B' \cap F| - r(\operatorname{span}(B' \cap F) \cap A)$. Hence

(43.60)
$$\begin{aligned}&\nu(E') \geq |B' \cap F| - r(\operatorname{span}(B' \cap F) \cap A)\\ &= |B'| - |B' \setminus F| - r(\operatorname{span}(B' \cap F) \cap A)\\ &= |B'| - \tfrac{1}{2}(r(B' \setminus F) + r(\operatorname{span}(B' \cap F) \cap A)) - \tfrac{1}{2} r(\operatorname{span}(B' \cap F) \cap A)\\ &\geq |B'| - \tfrac{1}{2} r(K_{\mathcal{B}}) - \tfrac{1}{2} r(A) > |B_0| - \tfrac{1}{2} r(K_{\mathcal{B}}) - \tfrac{1}{2} r(A)\\ &\geq |B_0| - \tfrac{1}{2}(r(B_0 \setminus F) + r(A)) - \tfrac{1}{2} r(A) = |B_0 \cap F| - r(A) = \nu'.\end{aligned}$$

The second inequality holds as $\operatorname{span}(B' \setminus F)$ and $\operatorname{span}(B' \cap F) \cap A$ are subspaces of $K_{\mathcal{B}}$ having intersection $\{\mathbf{0}\}$ (since B is a matching and as (43.56) holds). The last inequality follows from

(43.61)
$$\begin{aligned}&r(K_{\mathcal{B}}) = r(\operatorname{span}(B_0) \cap K_{\mathcal{B}})\\ &= r((\operatorname{span}(B_0 \setminus F) + \operatorname{span}(B_0 \cap F)) \cap K_{\mathcal{B}})\\ &= r(\operatorname{span}(B_0 \setminus F) + (\operatorname{span}(B_0 \cap F) \cap K_{\mathcal{B}}))\\ &= r(\operatorname{span}(B_0 \setminus F) + r(\operatorname{span}(B_0 \cap F) \cap K_{\mathcal{B}})).\end{aligned}$$

Here we use that

(43.62) $$(\mathrm{span}(B_0 \setminus F) + \mathrm{span}(B_0 \cap F)) \cap K_{\mathcal{B}} = \mathrm{span}(B_0 \setminus F) + (\mathrm{span}(B_0 \cap F) \cap K_{\mathcal{B}}),$$

which holds since if $x \in \mathrm{span}(B_0 \setminus F)$ and $y \in \mathrm{span}(B_0 \cap F)$ with $x+y \in K_{\mathcal{B}}$, then $y \in K_{\mathcal{B}}$ (since $x \in \mathrm{span}(B_0 \setminus F) \subseteq K_{\mathcal{B}}$ by (43.56)).

Now applying (43.54) repeatedly to $\mathcal{B}'$, we finally find a matching M' in E' with $|M'| = \nu' + 1$. It corresponds to a matching M in F with

(43.63) $$r(M \cup A) = 2|M| + r(A) = |M| + \nu' + 1 + r(A) = |B_0 \cap F| + |M| + 1.$$

Then applying (43.44) improves $\mathcal{B}$. ∎

The proof also yields an alternative proof of Theorem 43.2.

While most of the matroids we meet in daily life are linear, it might yet be interesting to extend the algorithm to the class of algebraic matroids. As Dress and Lovász [1987] remark, this requires the development of algorithmic techniques for algebraic matroids, for instance, for testing algebraic independence, and for finding a point p in the intersection of certain flats. If such techniques are available, pursuing the layout of the above algorithm for linear matroids might yield a polynomial-time algorithm for algebraic matroids.

An augmenting path algorithm for linear matroid matching, of complexity $O(n^3 m)$ (where $n :=$ rank, $m := |S|$) was given by Stallmann and Gabow [1984] and Gabow and Stallmann [1986] and an $O(n^4 m)$-time algorithm (by solving a sequence of matroid intersection algorithms) by Orlin and Vande Vate [1990] (these bounds can be improved to $O(n^{2.376} m)$ and $O(n^{3.376} m)$, respectively, with fast matrix multiplication).

43.9. Matroid matching is not polynomial-time solvable in general

Theorem 43.2 characterizes the matroid matching problem for algebraic matroids, and one is challenged to extend this to general matroids. A main objection to do this in a direct way is that in Theorem 43.2 a line of E may intersect the flat F in a point not contained in the original matroid. So we need to extend the matroid in some way, which is quite natural for linear matroids, but, as Lovász remarks, 'in general, there seems to be no hope to extend the original matroid so as to achieve the validity of [Theorem 43.2]. The possibility of "simulating" the flat F inside the matroid seems to be a difficult, and probably not only technical, question.'

Jensen and Korte [1982] and Lovász [1981] showed that, for matroids in general, the matroid matching problem is not solvable in polynomial time, if the matroid is given by an independence testing oracle (an oracle telling if a given set is independent or not). The construction in both papers is as follows.

Let $\nu \in \mathbb{Z}$, let S be a set, and let E be a partition of S into pairs. Let M be the matroid on S of rank 2ν, where $T \subseteq S$ is independent if and only if $|T| \leq 2\nu - 1$, or $|T| = 2\nu$ and T is not the union of ν pairs in E.

For each subset F of E of size ν, let M_F be the matroid on S obtained from M by adding $\bigcup F$ as independent subset.

It is easy to check that M and each of the M_F are matroids, and that E has no matroid matching of size ν with respect to M, while F is the unique matroid matching of size ν in M_F.

Suppose now that we want to find the maximum size of a matroid matching in a matroid, and that we know that the matroid is equal to M or to M_F for some ν-element $F \subseteq E$. Then we must ask the oracle for the independence of $\bigcup F$ for each ν-element subset F of E, in order to know if there exists a matroid matching of size ν. This takes exponential time.

This example shows that the matroid matching problem even does not belong to (oracle) co-NP, since any certificate that the matching number is at most $\nu - 1$, needs the oracle output that $\bigcup F$ is dependent, for all ν-element subsets F of E.

The example can be easily adapted to remove the oracle, and to obtain a proper problem in NP that is NP-complete. Let G be an undirected graph with vertex set V and let $\nu \in \mathbb{Z}_+$. For each vertex v of G, let p_v be a pair of elements, such that $p_u \cap p_v = \emptyset$ if $u \neq v$. Let $S := \bigcup_{v \in V} p_v$ and $E := \{p_v | v \in V\}$. So E is a partition of S into pairs. Define a matroid on S by extending the matroid M above by an independent set

$$I := \bigcup_{v \in C} p_v \tag{43.64}$$

for each clique C of G with $|C| = \nu$. Then E contains a matroid matching of size ν if and only if G has a clique of size ν. As the maximum-size clique problem is NP-complete, also the matroid matching problem for such matroids is NP-complete.

43.10. Further results and notes

43.10a. Optimal path-matching

Cunningham and Geelen [1996,1997] gave the following generalization of nonbipartite matching and matroid intersection.

Let $G = (V, E)$ be an undirected graph, let S_1 and S_2 be two disjoint stable subsets of V, and let $M_1 = (S_1, \mathcal{I}_1)$ and $M_2 = (S_2, \mathcal{I}_2)$ be matroids, with rank functions r_1 and r_2, such that $r_1(S_2) = r_2(S_2) =: \rho$. Define $R := V \setminus (S_1 \cup S_2)$. A *basic path-matching* is a collection of ρ vertex-disjoint $B_1 - B_2$ paths, each having all internal vertices in R, where B_1 and B_2 are bases of M_1 and M_2 respectively, together with a perfect matching on the vertices of R not covered by these paths.

If $R = V$, a basic path-matching is just a perfect matching. If $R = \emptyset$ and E consists of disjoint edges linking S_1 and S_2, then a basic path-matching corresponds to a common base.

Geelen and Cunningham showed that a basic path-matching exists if and only if for each $U_1 \subseteq S_1 \cup R$ and $U_2 \subseteq S_2 \cup R$ such that there is no edge connecting two sets among $U_1 \cap U_2$, $U_1 \setminus U_2$, $U_2 \setminus U_1$, one has

(43.65) $$r_1(S_1 \setminus U_1) + r_2(S_2 \setminus U_2) + |R \setminus (U_1 \cup U_2)| \geq \rho + o(G[U_1 \cap U_2]),$$

where $o(H)$ is the number of odd components of a graph H. Moreover, they gave a polynomial-time algorithm to decide whether there exists a basic path-matching.

More generally, they introduced the concept of an *independent path-matching*, which is a set F of edges such that each nonsingleton component of the graph (V, F) is an $S_1 \cup R - S_2 \cup R$ path all of whose internal vertices are in R, and such that the vertices in S_i covered by the paths is independent in M_i $(i = 1, 2)$. The corresponding *independent path-matching vector* is the vector $x \in \mathbb{Z}_+^E$ with $x(e) = 0$ if $e \notin F$, $x(e) = 2$ if $e \in F$ forms a component of (V, F) with both ends of e in R, and $x(e) = 1$ otherwise.

Geelen and Cunningham showed that the convex hull of the independent path-matching vectors is determined by:

(43.66)
$$\begin{array}{ll} x_e \geq 0 & \text{for } e \in E, \\ x(\delta(v)) \leq 2 & \text{for } v \in R, \\ x(E[U]) \leq |U \cap R| & \text{for } U \subseteq V \text{ with } U \cap S_1 = \emptyset \text{ or } U \cap S_2 = \emptyset, \\ x(E[U]) \leq |U| - 1 & \text{for } U \subseteq R, \\ x(\delta(U)) \leq r_i(U) & \text{for } U \subseteq S_i \text{ and } i = 1, 2, \end{array}$$

and that this system is TDI. It implies that the maximum of $\mathbf{1}^{\mathsf{T}}x$ over independent path-matching vectors is equal to the minimum of

(43.67) $$r_1(S_1 \setminus U_1) + r_2(S_2 \setminus U_2) + |R \setminus (U_1 \cup U_2)| + |R| - o(G[U_1 \cap U_2])$$

over all $U_i \subseteq S_i \cup R$ $(i = 1, 2)$ such that there is no edge connecting two sets among $U_1 \cap U_2$, $U_1 \setminus U_2$, $U_2 \setminus U_1$. (A simplified proof of this was given by Frank and Szegő [2002].)

Cunningham and Geelen argue that the set of inequalities (43.66) can be checked in polynomial time, implying (with the ellipsoid method) that, for any weight function w, an independent path-matching vector x maximizing $w^{\mathsf{T}}x$ can be found in strongly polynomial time. A combinatorial algorithm for the unweighted version was given by Spille and Weismantel [2002a,2002b].

For a survey, see Cunningham [2002].

43.10b. Further notes

Hochstättler and Kern [1989] showed that condition (43.13) is implied by the following:

(43.68) for any three flats A, B, C with
$$r(A \cup C) - r(A) = r(B \cup C) - r(B) = r(A \cup B \cup C) - r(A \cup B),$$
one has
$$r(\text{span}(A \cup C) \cap \text{span}(B \cup C)) - r(A \cap B) = r(A \cup C) - r(A).$$

Matroids with this property are called *pseudomodular* by Björner and Lovász [1987], who proved that full linear matroids (infinite matroids determined by linear independence of a linear space), full algebraic matroids (infinite matroids determined by algebraic independence of a field extension of a field), and full graphic matroids (cycle matroids of a complete graph) are pseudomodular. See also Lindström [1988], Dress, Hochstättler, and Kern [1994], and Tan [1997].

A randomized parallel algorithm for linear matroid matching was given by Narayanan, Saran, and Vazirani [1992,1994]. Stallmann and Gabow [1984] gave an algorithm for graphic matroid matching with running time $O(n^2m)$, which was improved by Gabow and Stallmann [1985] to $O(nm\log^6 n)$. Tong, Lawler, and Vazirani [1984] found a polynomial-time algorithm for *weighted* matroid matching for gammoids (by reduction to weighted matching). Structural properties of matroid matching, including an Edmonds-Gallai type decomposition, were given by Vande Vate [1992], which paper also studied the matroid matching polytope and a fractional relaxation of it.

The matroid matching problem generalizes the *matchoid problem* of J. Edmonds (cf. Jenkyns [1974]): given a graph $G = (V, E)$ and a matroid $M_v = (\delta(v), \mathcal{I}_v)$ for each v in V, what is the maximum number of edges such that the restriction to $\delta(v)$ forms an independent set in M_v, for each v in V?

Chapter 44

Submodular functions and polymatroids

In this chapter we describe some of the basic properties of a second main object of the present part, the submodular function. Each submodular function gives a polymatroid, which is a generalization of the independent set polytope of a matroid. We prove as a main result the theorem of Edmonds [1970b] that the vertices of a polymatroid are integer if and only if the associated submodular function is integer.

44.1. Submodular functions and polymatroids

Let f be a *set function* on a set S, that is, a function defined on the collection $\mathcal{P}(S)$ of all subsets of S. The function f is called *submodular* if

(44.1) $$f(T) + f(U) \geq f(T \cap U) + f(T \cup U)$$

for all subsets T, U of S. Similarly, f is called *supermodular* if $-f$ is submodular, i.e., if f satisfies (44.1) with the opposite inequality sign. f is *modular* if f is both submodular and supermodular, i.e., if f satisfies (44.1) with equality.

A set function f on S is called *nondecreasing* if $f(T) \leq f(U)$ whenever $T \subseteq U \subseteq S$, and *nonincreasing* if $f(T) \geq f(U)$ whenever $T \subseteq U \subseteq S$.

As usual, denote for each function $w : S \to \mathbb{R}$ and for each subset U of S,

(44.2) $$w(U) := \sum_{s \in U} w(s).$$

So w may be considered also as a set function on S, and one easily sees that w is modular, and that each modular set function f on S with $f(\emptyset) = 0$ may be obtained in this way. (More generally, each modular set function f on S satisfies $f(U) = w(U) + \gamma$ (for $U \subseteq S$), for some unique function $w : S \to \mathbb{R}$ and some unique real number γ.)

In a sense, submodularity is the discrete analogue of convexity. If we define, for any $f : \mathcal{P}(S) \to \mathbb{R}$ and any $x \in S$, a function $\delta f_x : \mathcal{P}(S) \to \mathbb{R}$ by: $\delta f_x(T) := f(T \cup \{x\}) - f(T)$, then f is submodular if and only if δf_x is nonincreasing for each $x \in S$.

In other words:

Theorem 44.1. *A set function f on S is submodular if and only if*

$$f(U \cup \{s\}) + f(U \cup \{t\}) \geq f(U) + f(U \cup \{s,t\}) \tag{44.3}$$

for each $U \subseteq S$ and distinct $s, t \in S \setminus U$.

Proof. Necessity being trivial, we show sufficiency. We prove (44.1) by induction on $|T \triangle U|$, the case $|T \triangle U| \leq 2$ being trivial (if $T \subseteq U$ or $U \subseteq T$) or being implied by (44.3). If $|T \triangle U| \geq 3$, we may assume by symmetry that $|T \setminus U| \geq 2$. Choose $t \in T \setminus U$. Then, by induction,

$$f(T \cup U) - f(T) \leq f((T \setminus \{t\}) \cup U) - f(T \setminus \{t\}) \leq f(U) - f(T \cap U), \tag{44.4}$$

(as $|T \triangle ((T \setminus \{t\}) \cup U)| < |T \triangle U|$ and $|(T \setminus \{t\}) \triangle U| < |T \triangle U|$). This shows (44.1). ∎

Define two polyhedra associated with a set function f on S:

$$\begin{aligned} P_f &:= \{x \in \mathbb{R}^S \mid x \geq \mathbf{0}, x(U) \leq f(U) \text{ for each } U \subseteq S\}, \\ EP_f &:= \{x \in \mathbb{R}^S \mid x(U) \leq f(U) \text{ for each } U \subseteq S\}. \end{aligned} \tag{44.5}$$

Note that P_f is nonempty if and only if $f \geq \mathbf{0}$, and that EP_f is nonempty if and only if $f(\emptyset) \geq 0$.

If f is a submodular function, then P_f is called the *polymatroid associated with* f, and EP_f the *extended polymatroid associated with* f. A polyhedron is called an (extended) polymatroid if it is the (extended) polymatroid associated with some submodular function. A polymatroid is bounded (since $0 \leq x_s \leq f(\{s\})$ for each $s \in S$), and hence is a polytope.

The following observation presents a basic technique in proofs for submodular functions, which we often use without further reference:

Theorem 44.2. *Let f be a submodular set function on S and let $x \in EP_f$. Then the collection of sets $U \subseteq S$ satisfying $x(U) = f(U)$ is closed under taking unions and intersections.*

Proof. Suppose $x(T) = f(T)$ and $x(U) = f(U)$. Then

$$\begin{aligned} f(T) + f(U) &\geq f(T \cap U) + f(T \cup U) \geq x(T \cap U) + x(T \cup U) \\ &= x(T) + x(U) = f(T) + f(U), \end{aligned} \tag{44.6}$$

implying that equality holds throughout. So $x(T \cap U) = f(T \cap U)$ and $x(T \cup U) = f(T \cup U)$. ∎

A vector x in EP_f (or in P_f) is called a *base vector* of EP_f (or of P_f) if $x(S) = f(S)$. A *base vector* of f is a base vector of EP_f. The set of all base vectors of f is called the *base polytope* of EP_f or of f. It is a face of EP_f, and denoted by B_f. So

$$B_f = \{x \in \mathbb{R}^S \mid x(U) \leq f(U) \text{ for all } U \subseteq S,\ x(S) = f(S)\}. \tag{44.7}$$

(It is a polytope, since $x_s = x(S) - x(S \setminus \{s\}) \geq f(S) - f(S \setminus \{s\})$ for each $s \in S$.)

Let f be a submodular set function on S and let $a \in \mathbb{R}^S$. Define the set function $f|a$ on S by

(44.8) $$(f|a)(U) := \min_{T \subseteq U}(f(T) + a(U \setminus T))$$

for $U \subseteq S$. It is easy to check that $f|a$ again is submodular and that

(44.9) $$EP_{f|a} = \{x \in EP_f \mid x \leq a\} \text{ and } P_{f|a} = \{x \in P_f \mid x \leq a\}.$$

It follows that if P is an (extended) polymatroid, then also the set $P \cap \{x \mid x \leq a\}$ is an (extended) polymatroid, for any vector a. In fact, as Lovász [1983c] observed, if $f(\emptyset) = 0$, then $f|a$ is the unique largest submodular function f' satisfying $f'(\emptyset) = 0$, $f' \leq f$, and $f'(U) \leq a(U)$ for each $U \subseteq V$.

44.1a. Examples

Matroids. Let $M = (S, \mathcal{I})$ be a matroid. Then the rank function r of M is submodular and nondecreasing. In Theorem 39.8 we saw that a set function r on S is the rank function of a matroid if and only if r is nonnegative, integer, nondecreasing and submodular with $r(U) \leq |U|$ for all $U \subseteq S$. (This last condition may be replaced by: $r(\emptyset) = 0$ and $r(\{s\}) \leq 1$ for each s in S.) Then the polymatroid P_r associated with r is equal to the independent set polytope of M (by Corollary 40.2b).

A generalization is obtained by partitioning S into sets $S_1, \ldots, S_k$, and defining

(44.10) $$f(J) := r(\bigcup_{i \in J} S_i)$$

for $J \subseteq \{1, \ldots, k\}$. It is not difficult to show that each integer nondecreasing submodular function f with $f(\emptyset) = 0$ can be constructed in this way (see Section 44.6b).

As another generalization, if $w : S \to \mathbb{R}_+$, define $f(U)$ to be the maximum of $w(I)$ over $I \in \mathcal{I}$ with $I \subseteq U$. Then f is submodular. (To see this, write $w = \lambda_1 \chi^{T_1} + \cdots + \lambda_n \chi^{T_n}$, with $\emptyset \neq T_1 \subset T_2 \subset \cdots \subset T_n \subseteq S$. Then by (40.3), $f(U) = \sum_{i=1}^n \lambda_i r(U \cap T_i)$, implying that f is submodular.)

For more on the relation between submodular functions and matroids, see Sections 44.6a and 44.6b.

Matroid intersection. Let $M_1 = (S, \mathcal{I}_1)$ and $M_2 = (S, \mathcal{I}_2)$ be matroids, with rank functions r_1 and r_2 respectively. Then the function f given by

(44.11) $$f(U) := r_1(U) + r_2(S \setminus U)$$

for $U \subseteq S$, is submodular. By the matroid intersection theorem (Theorem 41.1), the minimum value of f is equal to the maximum size of a common independent set.

Set unions. Let $T_1, \ldots, T_n$ be subsets of a finite set T and let $S = \{1, \ldots, n\}$. Define

(44.12) $$f(U) := |\bigcup_{i\in U} T_i|$$

for $U \subseteq S$. Then f is nondecreasing and submodular. More generally, for $w : T \to \mathbb{R}_+$, the function f defined by

(44.13) $$f(U) := w(\bigcup_{i\in U} T_i)$$

for $U \subseteq S$, is nondecreasing and submodular.

More generally, for any nondecreasing submodular set function g on T, the function f defined by

(44.14) $$f(U) := g(\bigcup_{i\in U} T_i)$$

for $U \subseteq S$, again is nondecreasing and submodular.

Let $G = (V, E)$ be the bipartite graph corresponding to $T_1, \ldots, T_n$. That is, G has colour classes S and T, and $s \in S$ and $t \in T$ are adjacent if and only if $t \in T_s$. Then we have: $x \in P_f$ if and only if there exist $z \in P_g$ and $y : E \to \mathbb{Z}_+$ such that

(44.15) $$\begin{aligned} y(\delta(v)) &= x(v) \quad \text{for all } v \in S,\\ y(\delta(v)) &= z(v) \quad \text{for all } v \in T. \end{aligned}$$

So y may be considered as an 'assignment' of a 'supply' z to a 'demand' x. If g and x are integer we can take also y and z integer.

Directed graph cut functions. Let $D = (V, A)$ be a directed graph and let $c : A \to \mathbb{R}_+$ be a 'capacity' function on A. Define

(44.16) $$f(U) := c(\delta^{\text{out}}(U))$$

for $U \subseteq V$ (where $\delta^{\text{out}}(U)$ denotes the set of arcs leaving U). Then f is submodular (but in general not nondecreasing). A function f arising in this way is called a *cut function*.

Hypergraph cut functions. Let $(V, \mathcal{E})$ be a hypergraph. For $U \subseteq V$, let $f(U)$ be the number of edges $E \in \mathcal{E}$ split by U (that is, with both $E \cap U$ and $E \setminus U$ nonempty). Then f is submodular.

Directed hypergraph cut functions. Let V be a finite set and let $(E_1, F_1), \ldots, (E_m, F_m)$ be pairs of subsets of V. For $U \subseteq V$, let $f(U)$ be the number of indices i with $U \cap E_i \neq \emptyset$ and $F_i \not\subseteq U$. Then f is submodular. (In proving this, we can assume $m = 1$, since any sum of submodular functions is submodular again.)

More generally, we can choose $c_1, \ldots, c_m \in \mathbb{R}_+$ and define

(44.17) $$f(U) = \sum (c_i \mid U \cap E_i \neq \emptyset, F_i \not\subseteq U)$$

for $U \subseteq V$. Again, f is submodular. This generalizes the previous two examples (where $E_i = F_i$ for each i or $|E_i| = |F_i| = 1$ for each i).

Maximal element. Let V be a finite set and let $h : V \to \mathbb{R}$. For nonempty $U \subseteq V$, define

(44.18) $$f(U) := \max\{h(u) \mid u \in U\},$$

and define $f(\emptyset)$ to be the minimum of $h(v)$ over $v \in V$. Then f is submodular.

Subtree diameter. Let $G = (V, E)$ be a forest (a graph without circuits), and for each $X \subseteq E$ define

(44.19) $$f(X) := \sum_K \text{diameter}(K),$$

where K ranges over the components of the graph (V, X). Here diameter(K) is the length of a longest path in K. Then f is submodular (Tamir [1993]); that is:

(44.20) $$f(X) + f(Y) \geq f(X \cap Y) + f(X \cup Y)$$

for $X, Y \subseteq E$.

To see this, denote, for any $X \subseteq E$, the set of vertices covered by X by VX. We first show (44.20) for $X, Y \subseteq E$ with (VX, X) and (VY, Y) connected and $VX \cap VY \neq \emptyset$. Note that in this case $X \cap Y$ and $X \cup Y$ give connected subgraphs again.

The proof of (44.20) is based on the fact that for all $s, t, u, v \in V$ one has:

(44.21) $$\begin{aligned}&\text{dist}(s, u) + \text{dist}(t, v) \geq \text{dist}(s, t) + \text{dist}(u, v)\\ &\text{or } \text{dist}(t, u) + \text{dist}(s, v) \geq \text{dist}(s, t) + \text{dist}(u, v),\end{aligned}$$

where dist denotes the distance in G.

To prove (44.20), let P and Q be longest paths in $X \cap Y$ and $X \cup Y$ respectively. If EQ is contained in X or in Y, then (44.20) follows, since P is contained in X and in Y. So we can assume that EQ is contained neither in X nor in Y. Let Q have ends u, v, with $u \in VX$ and $v \in VY$. Let P have ends s, t. So $s, t, u \in VX$ and $s, t, v \in VY$. Hence (44.21) implies (44.20).

We now derive (44.20) for all $X, Y \subseteq E$. Let $\mathcal{X}$ and $\mathcal{Y}$ be the collections of edge sets of the components of (V, X) and of (V, Y) respectively. Let $\mathcal{F}$ be the family made by the union of $\mathcal{X}$ and $\mathcal{Y}$, taking the sets in $\mathcal{X} \cap \mathcal{Y}$ twice. Then

(44.22) $$f(X) + f(Y) \geq \sum_{Z \in \mathcal{F}} f(Z).$$

We now modify $\mathcal{F}$ iteratively as follows. If $Z, Z' \in \mathcal{F}$, $Z \not\subseteq Z' \not\subseteq Z$, and $VZ \cap VZ' \neq \emptyset$, we replace Z, Z' by $Z \cap Z'$ and $Z \cup Z'$. By (44.20), (44.22) is maintained. By Theorem 2.1, these iterations stop. We delete the empty sets in the final $\mathcal{F}$.

Then the inclusionwise maximal sets in $\mathcal{F}$ have union equal to $X \cup Y$ and form the nonempty edge sets of the components of $(V, X \cup Y)$. Similarly, the inclusionwise minimal sets in $\mathcal{F}$ form the nonempty edge sets of the components of $(V, X \cap Y)$. So

(44.23) $$\sum_{Z \in \mathcal{F}} f(Z) = f(X \cap Y) + f(X \cup Y),$$

and we have (44.20).

Further examples. Choquet [1951,1955] showed that the classical Newtonian capacity in $\mathbb{R}^3$ is submodular. Examples of submodular functions based on probability are given by Fujishige [1978b] and Han [1979], and other examples by Lovász [1983c].

44.2. Optimization over polymatroids by the greedy method

Edmonds [1970b] showed that one can optimize a linear function $w^{\mathsf{T}}x$ over an (extended) polymatroid by an extension of the greedy algorithm. The submodular set function f on S is given by a *value giving oracle*, that is, by an oracle that returns $f(U)$ for any $U \subseteq S$.

Let f be a submodular set function on S, and suppose that we want to maximize $w^{\mathsf{T}}x$ over EP_f, for some $w : S \to \mathbb{R}$. We can assume that $EP_f \neq \emptyset$, that is $f(\emptyset) \geq 0$, and hence that $f(\emptyset) = 0$ (since decreasing $f(\emptyset)$ maintains submodularity). We can also assume that $w \geq \mathbf{0}$, since if some component of w is negative, the maximum value is unbounded.

Now order the elements in S as $s_1, \ldots, s_n$ such that $w(s_1) \geq \cdots \geq w(s_n)$. Define

(44.24) $\qquad U_i := \{s_1, \ldots, s_i\}$ for $i = 0, \ldots, n$,

and define $x \in \mathbb{R}^S$ by

(44.25) $\qquad x(s_i) := f(U_i) - f(U_{i-1})$ for $i = 1, \ldots, n$.

Then x maximizes $w^{\mathsf{T}}x$ over EP_f, as will be shown in the following theorem.

To prove it, consider the following linear programming duality equation:

(44.26)
$$\max\{w^{\mathsf{T}}x \mid x \in EP_f\} = \min\{\sum_{T \subseteq S} y(T)f(T) | y \in \mathbb{R}_+^{\mathcal{P}(S)}, \sum_{T \in \mathcal{P}(S)} y(T)\chi^T = w\}.$$

Define:

(44.27)
$$\begin{aligned} y(U_i) &:= w(s_i) - w(s_{i+1}) && (i = 1, \ldots, n-1),\\ y(S) &:= w(s_n), &&\\ y(T) &:= 0 && (T \neq U_i \text{ for each } i). \end{aligned}$$

Theorem 44.3. *Let f be a submodular set function on S with $f(\emptyset) = 0$ and let $w : S \to \mathbb{R}_+$. Then x and y given by* (44.25) *and* (44.27) *are optimum solutions of* (44.26).

Proof. We first show that x belongs to EP_f; that is, $x(T) \leq f(T)$ for each $T \subseteq S$. This is shown by induction on $|T|$, the case $T = \emptyset$ being trivial. Let $T \neq \emptyset$ and let k be the largest index with $s_k \in T$. Then by induction,

(44.28) $\qquad x(T \setminus \{s_k\}) \leq f(T \setminus \{s_k\})$.

Hence

(44.29) $\qquad x(T) \leq f(T \setminus \{s_k\}) + x(s_k) = f(T \setminus \{s_k\}) + f(U_k) - f(U_{k-1}) \leq f(T)$

(the last inequality follows from the submodularity of f). So $x \in EP_f$.

Also, y is feasible for (44.26). Trivially, $y \geq \mathbf{0}$. Moreover, for any i we have by (44.27):

$$\sum_{T \ni s_i} y(T) = \sum_{j \geq i} y(U_j) = w(s_i). \tag{44.30}$$

So y is a feasible solution of (44.26).

Optimality of x and y follows from:

$$\begin{aligned} w^{\mathsf{T}}x &= \sum_{s \in S} w(s)x_s = \sum_{i=1}^{n} w(s_i)(f(U_i) - f(U_{i-1})) \\ &= \sum_{i=1}^{n-1} f(U_i)(w(s_i) - w(s_{i+1})) + f(S)w(s_n) = \sum_{T \subseteq S} y(T)f(T). \end{aligned} \tag{44.31}$$

The third equality follows from a straightforward reordering of the terms, using that $f(\emptyset) = 0$. ∎

Note that if f is integer, then x is integer, and that if w is integer, then y is integer. Moreover, if f is nondecreasing, then x is nonnegative. Hence, in that case, x and y are optimum solutions of

$$\begin{aligned} &\max\{w^{\mathsf{T}}x \mid x \in P_f\} \\ &= \min\{\sum_{T \subseteq S} y(T)f(T) \mid y \in \mathbb{R}_+^{\mathcal{P}(S)}, \sum_{T \in \mathcal{P}(S)} y(T)\chi^T \geq w\}. \end{aligned} \tag{44.32}$$

Therefore:

Corollary 44.3a. *Let f be a nondecreasing submodular set function on S with $f(\emptyset) = 0$ and let $w : S \to \mathbb{R}_+$. Then x and y given by* (44.25) *and* (44.27) *are optimum solutions for* (44.32).

Proof. Directly from Theorem 44.3, using the fact that $x \geq \mathbf{0}$ if f is nondecreasing. ∎

As for complexity we have:

Corollary 44.3b. *Given a submodular set function f on a set S (by a value giving oracle) and a function $w \in \mathbb{Q}^S$, we can find an $x \in EP_f$ maximizing $w^{\mathsf{T}}x$ in strongly polynomial time. If f is moreover nondecreasing, then $x \in P_f$ (and hence x maximizes $w^{\mathsf{T}}x$ over P_f).*

Proof. By the extension of the greedy method given above. ∎

The greedy algorithm can be interpreted geometrically as follows. Let w be some linear objective function on S, with $w(s_1) \geq \ldots \geq w(s_n)$. Travel via the vertices of P_f along the edges of P_f, by starting at the origin, as follows: first go from the origin as far as possible (in P_f) in the positive s_1-direction, say to vertex x_1; next go from x_1 as far as possible in the positive s_2-direction, say to x_2, and so on. After n steps one reaches a vertex x_n maximizing $w^{\mathsf{T}}x$

over P_f. In fact, the effectiveness of this algorithm characterizes polymatroids (Dunstan and Welsh [1973]).

44.3. Total dual integrality

Theorem 44.3 implies the box-total dual integrality of the following system:

(44.33) $\quad x(U) \leq f(U)$ for $U \subseteq S$.

Corollary 44.3c. *If f is submodular, then* (44.33) *is box-totally dual integral.*

Proof. Consider the dual of maximizing $w^{\mathsf{T}}x$ over (44.33), for some $w \in \mathbb{Z}_+^S$. By Theorem 44.3, it has an optimum solution $y : \mathcal{P}(S) \to \mathbb{R}_+$ with the sets $U \subseteq S$ having $y(U) > 0$ forming a chain. So these constraints give a totally unimodular submatrix of the constraint matrix (by Theorem 41.11). Therefore, by Theorem 5.35, (44.33) is box-TDI. ∎

This gives the integrality of polyhedra:

Corollary 44.3d. *For any integer submodular set function f, the polymatroid P_f and the extended polymatroid EP_f are integer.*

Proof. Directly from Corollary 44.3c. (In fact, integer optimum solutions are explicitly given by Theorem 44.3 and Corollary 44.3a.) ∎

44.4. f is determined by EP_f

Theorem 44.3 implies that for any extended polymatroid P there is a unique submodular function f satisfying $f(\emptyset) = 0$ and $EP_f = P$, since:

Corollary 44.3e. *Let f be a submodular set function on S with $f(\emptyset) = 0$. Then*

(44.34) $\quad f(U) = \max\{x(U) \mid x \in EP_f\}$

for each $U \subseteq S$.

Proof. Directly from Theorem 44.3 by taking $w := \chi^U$. ∎

So there is a one-to-one correspondence between nonempty extended polymatroids and submodular set functions f with $f(\emptyset) = 0$. The correspondence relates integer extended polymatroids with integer submodular functions.

There is a similar correspondence between nonempty polymatroids and *nondecreasing* submodular set functions f with $f(\emptyset) = 0$. For any (not necessarily nondecreasing) nonnegative submodular set function f, define $\bar{f}$ by:

(44.35) $\bar{f}(\emptyset) = 0,$
$$\bar{f}(U) = \min_{T \supseteq U} f(T) \quad \text{for nonempty } U \subseteq S.$$

It is easy to see that $\bar{f}$ is nondecreasing and submodular and that $P_{\bar{f}} = P_f$ (Dunstan [1973]). In fact, $\bar{f}$ is the unique nondecreasing submodular set function associated with P_f, with $\bar{f}(\emptyset) = 0$, as (Kelley [1959]):

Corollary 44.3f. *If f is a nondecreasing submodular function with $f(\emptyset) = 0$, then*

(44.36) $$f(U) = \max\{x(U) \mid x \in P_f\}$$

for each $U \subseteq S$.

Proof. This follows from Corollary 44.3a by taking $w := \chi^T$. ∎

This one-to-one correspondence between polymatroids and nondecreasing submodular set functions f with $f(\emptyset) = 0$ relates integer polymatroids to integer such functions:

Corollary 44.3g. *For each integer polymatroid P there exists a unique integer nondecreasing submodular function f with $f(\emptyset) = 0$ and $P = P_f$.*

Proof. By Corollary 44.3d and (44.36). ∎

By (44.36) we have for any nonnegative submodular set function f that $\bar{f}(U) = \max\{x(U) \mid x \in P_f\}$. Since we can optimize over EP_f in polynomial time (with the greedy algorithm described above), with the ellipsoid method we can optimize over $P_f = EP_f \cap \mathbb{R}^S_+$ in polynomial time. Hence we can calculate $\bar{f}(U)$ in polynomial time. Alternatively, calculating $\bar{f}(U)$ amounts to minimizing the submodular function $f'(T) := f(T \cup U)$.

In fact $\bar{f}$ is the largest among all nondecreasing submodular set functions g on S with $g(\emptyset) = 0$ and $g \leq f$, as can be checked straightforwardly.

44.5. Supermodular functions and contrapolymatroids

Similar results hold for supermodular functions and the associated contrapolymatroids. Associate the following polyhedra with a set function g on S:

(44.37) $$Q_g := \{x \in \mathbb{R}^S \mid x \geq \mathbf{0}, x(U) \geq g(U) \text{ for each } U \subseteq S\},$$
$$EQ_g := \{x \in \mathbb{R}^S \mid x(U) \geq g(U) \text{ for each } U \subseteq S\}.$$

If g is supermodular, then Q_g and EQ_g are called the *contrapolymatroid* and the *extended contrapolymatroid associated with* g, respectively. A vector $x \in EQ_g$ (or Q_g) is called a *base vector* of EQ_g (or Q_g) if $x(S) = g(S)$. A *base vector* of g is a base vector of EQ_g.

Since $EQ_g = -EP_{-g}$, we can reduce most problems on (extended) contrapolymatroids to (extended) polymatroids. Again we can minimize a linear function $w^{\mathsf{T}}x$ over EQ_g with the greedy algorithm, as described in Section 44.2. (In fact, we can apply the same formulas (44.25) and (44.27) for g instead of f.) If g is nondecreasing, it yields a nonnegative optimum solution, and hence a vector x minimizing $w^{\mathsf{T}}x$ over Q_g.

Similarly, the system

(44.38) $\quad x(U) \geq g(U)$ for $U \subseteq S$

is box-TDI, as follows directly from the box-total dual integrality of

(44.39) $\quad x(U) \leq -g(U)$ for $U \subseteq S$.

Let EP_f be the extended polymatroid associated with the submodular function f with $f(\emptyset) = 0$. Let B_f be the face of base vectors of EP_f, i.e.,

(44.40) $\quad B_f = \{x \in EP_f \mid x(S) = f(S)\}$.

A vector $y \in \mathbb{R}^S$ is called *spanning* if there exists an x in B_f with $x \leq y$. Let Q be the set of spanning vectors.

A vector y belongs to Q if and only if $(f|y)(S) = f(S)$, that is (by (44.8) and (44.9)) if and only if

(44.41) $\quad y(U) \geq f(S) - f(S \setminus U)$

for each $U \subseteq S$. So Q is equal to the contrapolymatroid EQ_g associated with the submodular function g defined by $g(U) := f(S) - f(S \setminus U)$ for $U \subseteq S$. Then B_f is equal to the face of minimal elements of EQ_g.

There is a one-to-one correspondence between submodular set functions f on S with $f(\emptyset) = 0$ and supermodular set functions g on S with $g(\emptyset) = 0$, given by the relations

(44.42) $\quad g(U) = f(S) - f(S \setminus U)$ and $f(U) = g(S) - g(S \setminus U)$

for $U \subseteq S$.

Then the pair $(-g, -Q)$ is related to the pair (f, P) by a relation similar to the duality relation of matroids (cf. Section 44.6f).

44.6. Further results and notes

44.6a. Submodular functions and matroids

Let P be the polymatroid associated with the nondecreasing integer submodular set function f on S, with $f(\emptyset) = 0$. Then the collection

(44.43) $\quad \mathcal{I} := \{I \subseteq S \mid \chi^I \in P\}$

forms the collection of independent sets of a matroid $M = (S, \mathcal{I})$ (this result was announced by Edmonds and Rota [1966] and proved by Pym and Perfect [1970]). By Corollary 40.2b, the subpolymatroid (cf. Section 44.6c)

(44.44) $$P|\mathbf{1} = \{x \in P \mid x \leq \mathbf{1}\}$$

is the convex hull of the incidence vectors of the independent sets of M. By (44.8), the rank function r of M satisfies

(44.45) $$r(U) = \min_{T \subseteq U}(|U \setminus T| + f(T))$$

for $U \subseteq S$.

As an example, if f is the submodular function given in the set union example in Section 44.1a, we obtain the transversal matroid on $\{1,\ldots,n\}$ with $I \subseteq \{1,\ldots,n\}$ independent if and only if the family $(T_i \mid i \in I)$ has a transversal (Edmonds [1970b]).

44.6b. Reducing integer polymatroids to matroids

In fact, each integer polymatroid can be derived from a matroid as follows (Helgason [1974]). Let f be a nondecreasing submodular set function on S with $f(\emptyset) = 0$. Choose for each s in S, a set X_s of size $f(\{s\})$, such that the sets X_s ($s \in S$) are disjoint. Let $X := \bigcup_{s\in S} X_s$, and define a set function r on X by

(44.46) $$r(U) := \min_{T \subseteq S}(|U \setminus \bigcup_{s\in T} X_s| + f(T))$$

for $U \subseteq X$. One easily checks that r is the rank function of a matroid M (by checking the axioms (39.38)), and that for each subset T of S

(44.47) $$f(T) = r(\bigcup_{s\in T} X_s).$$

Therefore, f arises from the rank function of M, as in the Matroids example in Section 44.1a. The polymatroid P_f associated with f is just the convex hull of all vectors x for which there exists an independent set I in M with $x_s = |I \cap X_s|$ for all s in S.

Given a nondecreasing submodular set function f on S with $f(\emptyset) = 0$, Lovász [1980a] called a subset $U \subseteq S$ a *matching* if

(44.48) $$f(U) = \sum_{s\in U} f(\{s\}).$$

If $f(\{s\}) = 1$ for each s in S, f is the rank function of a matroid, and U is a matching if and only if U is independent in this matroid. If $f(\{s\}) = 2$ for each s in S, the elements of S correspond to certain flats of rank 2 in a matroid. Now determining the maximum size of a matching is just the matroid matching problem (cf. Chapter 43).

44.6c. The structure of polymatroids

Vertices of polymatroids (Edmonds [1970b], Shapley [1965,1971]). Let f be a submodular set function on a set $S = \{s_1,\ldots,s_n\}$ with $f(\emptyset) = 0$. Let P_f be the polymatroid associated with f. It follows immediately from the greedy algorithm, as in the proof of Corollary 44.3a, that the vertices of P_f are given by (for $i = 1,\ldots,n$):

(44.49) $$x(s_{\pi(i)}) = \begin{cases} f(\{s_{\pi(1)}, \ldots, s_{\pi(i)}\}) - f(\{s_{\pi(1)}, \ldots, s_{\pi(i-1)}\}) & \text{if } i \leq k, \\ 0 & \text{if } i > k, \end{cases}$$

where π ranges over all permutations of $\{1, \ldots, n\}$ and where k ranges over $0, \ldots, n$.

Similarly, for any submodular set function f on S with $f(\emptyset) = 0$, the vertices of the extended polymatroid EP_f are given by

(44.50) $$x(s_{\pi(i)}) = f(\{s_{\pi(1)}, \ldots, s_{\pi(i)}\}) - f(\{s_{\pi(1)}, \ldots, s_{\pi(i-1)}\})$$

for $i = 1, \ldots, n$, where π ranges over all permutations of $\{1, \ldots, n\}$.

Topkis [1984] characterized adjacency of the vertices of a polymatroid, while Bixby, Cunningham, and Topkis [1985] and Topkis [1992] gave further results on vertices of and paths on a polymatroid and on related partial orders of S.

Facets of polymatroids. Let f be a nondecreasing submodular set function on S with $f(\emptyset) = 0$. One easily checks that P_f is full-dimensional if and only if $f(\{s\}) > 0$ for all s in S. If P_f is full-dimensional there is a unique minimal collection of linear inequalities defining P_f (clearly, up to scalar multiplication). They correspond to the facets of P_f. Edmonds [1970b] found that this collection is given by the following theorem. A subset $U \subseteq S$ is called an *f-flat* if $f(U \cup \{s\}) > f(U)$ for all $s \in S \setminus U$, and U is called *f-inseparable* if there is no partition of U into nonempty sets U_1 and U_2 with $f(U) = f(U_1) + f(U_2)$. Then:

Theorem 44.4. *Let f be a nondecreasing submodular set function on S with $f(\emptyset) = 0$ and $f(\{s\}) > 0$ for each $s \in S$. The following is a minimal system determining the polymatroid P_f:*

(44.51) $$\begin{array}{ll} x_s \geq 0 & (s \in S), \\ x(U) \leq f(U) & (U \text{ is a nonempty } f\text{-inseparable } f\text{-flat}). \end{array}$$

Proof. It is easy to see that (44.51) determines P_f, as any other inequality $x(U) \leq f(U)$ follows from (44.51). The irredundancy of collection (44.51) can be seen as follows.

Clearly, each inequality $x_s \geq 0$ determines a facet. Next consider a nonempty f-inseparable f-flat U. Suppose that the face determined by U is not a facet. Then it is contained in another face, say determined by T. Let x be a vertex of P_f with $x(U \setminus T) = f(U \setminus T)$, $x(U) = f(U)$, and $x(S \setminus U) = 0$. Such a vertex exists by the greedy algorithm (cf. (44.49)).

Since x is on the face determined by U, it is also on the face determined by T. So $x(T) = f(T)$. Hence $f(T) = x(T) = x(T \cap U) = f(U) - f(U \setminus T)$. So we have equality throughout in:

(44.52) $$f(U \setminus T) + f(T) \geq f(U \setminus T) + f(T \cap U) \geq f(U).$$

This implies that $U \setminus T = \emptyset$ or $T \cap U = \emptyset$ (as U is f-inseparable), and that $f(T) = f(T \cap U)$. If $U \setminus T = \emptyset$, then $U \subset T$, and hence (as U is an f-flat) $f(T) > f(U) \geq f(T \cap U)$, a contradiction. If $T \cap U = \emptyset$, then $f(T) = f(T \cap U) = 0$, implying that $T = \emptyset$, again a contradiction. ∎

It follows that the face $\{x \in P_f \mid x(S) = f(S)\}$ of maximal vectors in P_f is a facet if and only if $f(U) + f(S \setminus U) > f(S)$ for each proper nonempty subset U of S. More generally, its codimension is equal to the number of inclusionwise minimal nonempty sets U with $f(U) + f(S \setminus U) = f(S)$ (cf. Fujishige [1984a]).

Faces of polymatroids (Giles [1975]). We now extend the characterizations of vertices and facets of polymatroids given above to arbitrary faces. Let P be the polymatroid associated with the nondecreasing submodular set function f on S with $f(\emptyset) = 0$. Suppose that P is full-dimensional. If $\emptyset \neq S_1 \subset \cdots \subset S_k \subseteq T \subseteq S$, then

(44.53) $$F = \{x \in P \mid x(S_1) = f(S_1), \ldots, x(S_k) = f(S_k), x(S \setminus T) = 0\}$$

is a face of P of dimension at most $|T| - k$. (Indeed, F is nonempty by the characterization (44.49) of vertices, while $\dim(F) \leq |T| - k$, as the incidence vectors of $S_1, \ldots, S_k$ are linearly independent.)

In fact, each face has a representation (44.53). Indeed, let F be a face of P. Define $T = \{s \in S \mid x_s > 0 \text{ for some } x \text{ in } F\}$, and let $S_1 \subset \cdots \subset S_k$ be any maximal chain of nonempty subsets of T with the property that

(44.54) $$F \subseteq \{x \in P \mid x(S_1) = f(S_1), \ldots, x(S_k) = f(S_k), x(S \setminus T) = 0\}.$$

Then we have equality in (44.54), and $\dim(F) = |T| - k$. (Here a maximal chain is a chain which is contained in no larger chain satisfying (44.54) — since the empty chain satisfies (44.54), there exist maximal chains.)

In order to prove this assertion, suppose that F has dimension d. As the right-hand side of (44.54) is a face of P of dimension at most $|T| - k$, it suffices to show that $d = |T| - k$. Therefore, suppose $d < |T| - k$. Then there exists a subset U of S such that $x(U) = f(U)$ for all x in F, and such that the incidence vector of $U \cap T$ is linearly independent of the incidence vectors of $S_1, \ldots, S_k$. That is, $U \cap T$ is not the union of some of the sets $S_i \setminus S_{i-1}$ ($i = 1, \ldots, k$). Since $x(U \cap T) = x(U) = f(U) \geq f(U \cap T)$ for all x in F, we may assume that $U \subseteq T$. Since the collection of subsets U of S with $x(U) = f(U)$ is closed under taking unions and intersections, we may assume moreover that U is comparable with each of the sets in the chain $S_1 \subset \cdots \subset S_k$. Hence U could be added to the chain to obtain a larger chain, contradicting our assumption. So $d = |T| - k$.

Note that a chain $S_1 \subset \cdots \subset S_k$ of nonempty subsets of T is a maximal chain satisfying (44.54) if and only if there is equality in (44.54) and (setting $S_0 := \emptyset$):

(44.55) $f(S_k \cup \{s\}) > f(S_k)$ for all s in $T \setminus S_k$, and each of the sets $S_i \setminus S_{i-1}$ is f_i-inseparable, where f_i is the submodular set function on $S_i \setminus S_{i-1}$ given by $f_i(U) := f(U \cup S_{i-1}) - f(S_{i-1})$ for $U \subseteq S_i \setminus S_{i-1}$.

This may be derived straightforwardly from the existence, by (44.49), of appropriate vertices of F.

It is not difficult to show that if F has a representation (44.53), then F is the direct sum of $F_1, \ldots, F_k$ and Q, where F_i is the face of maximal vectors in the polymatroid associated with f_i ($i = 1, \ldots, k$), and Q is the polymatroid associated with the submodular set function g on $T \setminus S_k$ given by $g(U) := f(U \cup S_k) - f(S_k)$ for $U \subseteq T \setminus S_k$. Since $\dim(F_i) \leq |S_i \setminus S_{i-1}| - 1$ and $\dim(Q) \leq |T \setminus S_k|$, this yields that $\dim(F) = |T| - k$ if and only if $\dim(F_i) = |S_i \setminus S_{i-1}| - 1$ ($i = 1, \ldots, k$) and $\dim(Q) = |T \setminus S_k|$. From this, characterization (44.55) can be derived again. It also yields that if F, represented by (44.53), has dimension $|T| - k$, then the unordered partition $\{S_1, S_2 \setminus S_1, \ldots, S_k \setminus S_{k-1}, T \setminus S_k\}$ is the same for all maximal chains $S_1 \subset \cdots \subset S_k$.

For a characterization of the faces of a polymatroid, see Fujishige [1984a].

44.6d. Characterization of polymatroids

Let P be the polymatroid associated with the nondecreasing submodular set function f on S with $f(\emptyset) = 0$. The following three observations are easily derived from the representation (44.49) of vertices of P. (a) If x_0 is a vertex of P, there exists a vertex x_1 of P such that $x_1 \geq x_0$ and x_1 has the form (44.49) with $k = n$. (b) A vertex x_1 of P can be represented as (44.49) with $k = n$ if and only if $x_1(S) = f(S)$. (c) The convex hull of the vertices x_1 of P with $x_1(S) = f(S)$ is the face $\{x \in P \mid x(S) = f(S)\}$ of P. It follows directly from (a), (b) and (c) that $x \in P$ is a maximal element of P (with respect to $\leq$) if and only if $x(S) = f(S)$. So for each vector y in P there is a vector x in P with $y \leq x$ and $x(S) = f(S)$.

Applying this to the subpolymatroids $P|a = P \cap \{x \mid x \leq a\}$ (cf. Section 44.1), one finds the following property of polymatroids:

(44.56) for each $a \in \mathbb{R}^S_+$ there exists a number $r(a)$ such that each maximal vector x of $P \cap \{x \mid x \leq a\}$ satisfies $x(S) = r(a)$.

Here *maximal* is maximal in the partial order $\leq$ on vectors. The number $r(a)$ is called the *rank* of a, and any x with the properties mentioned in (44.56) is called a *base* of a.

Edmonds [1970b] (cf. Dunstan [1973], Woodall [1974b]) noticed the following (we follow the proof of Welsh [1976]):

Theorem 44.5. *Let $P \subseteq \mathbb{R}^S_+$. Then P is a polymatroid if and only if P is compact, and satisfies* (44.56) *and*

(44.57) *if $\mathbf{0} \leq y \leq x \in P$, then $y \in P$.*

Proof. Necessity was observed above. To see sufficiency, let f be the set function on S defined by

(44.58) $f(U) := \max\{x(U) \mid x \in P\}$

for $U \subseteq S$. Then f is nonnegative and nondecreasing. Moreover, f is submodular. To see this, consider $T, U \subseteq S$. Let x be a maximal vector in P satisfying $x_s = 0$ if $s \notin T \cup U$, and let y be a maximal vector in P satisfying $y(s) = 0$ if $s \notin T \cap U$ and $x \leq y$. Note that (44.56) and (44.58) imply that $x(T \cap U) = f(T \cap U)$ and $y(T \cup U) = f(T \cup U)$. Hence

(44.59) $$\begin{aligned} f(T)+f(U) &\geq y(T)+y(U) = y(T\cap U)+y(T\cup U) \geq x(T\cap U)+y(T\cup U) \\ &= f(T \cap U) + f(T \cup U), \end{aligned}$$

that is, f is submodular.

We finally show that P is equal to the polymatroid P_f associated to f. Clearly, $P \subseteq P_f$, since if $x \in P$ then $x(U) \leq f(U)$ for each $U \subseteq S$, by definition (44.58) of f.

To see that $P_f = P$, suppose $v \in P_f \setminus P$. Let u be a base of v (that is, a maximal vector $u \in P$ satisfying $u \leq v$). Choose u such that the set

(44.60) $U := \{s \in S \mid u_s < v_s\}$

is as large as possible. Since $v \notin P$, we have $u \neq v$, and hence $U \neq \emptyset$. As $v \in P_f$, we know

(44.61) $u(U) < v(U) \leq f(U)$.

Define

(44.62) $w := \frac{1}{2}(u + v)$.

So $u \leq w \leq v$. Hence u is a base of w, and each base of w is a base of v.

For any $z \in \mathbb{R}^S$, define z' as the projection of z on the subspace $L := \{x \in \mathbb{R}^S \mid x_s = 0 \text{ if } s \in S \setminus U\}$. That is:

(44.63) $z'(s) := z(s)$ if $s \in U$, and $z'(s) := 0$ if $s \in S \setminus U$.

By definition of f, there is an $x \in P$ with $x(U) = f(U)$. We may assume that $x \in L$. Choose $y \in L$ with $x \leq y$ and $u' \leq y$. Then

(44.64) $x(S) = x(U) = f(U) > u(U) = u'(U) = u'(S)$.

So $r(y) > u'(S)$. Hence, by (44.56), there exists a base z of y with $u' \leq z$ and $z(S) > u'(S)$. So $u'_s < z_s$ for at least one $s \in U$. This implies, since $u'_s < w'_s$ for each $s \in S$, that there is an $a \in P$ with $u' \leq a \leq w'$ and $a \neq u'$, hence $a(U) > u'(U)$.

Since $a \leq w' \leq w$, there is a base b of w with $a \leq b$. Then $b(S) = u(S)$ (since also u is a base of w) and $b(U) \geq a(U) > u'(U) = u(U)$. Hence $b_s < u_s = v_s$ for some $s \in S \setminus U$. Moreover, $b_s \leq w_s < v_s$ for each $s \in U$. So U is properly contained in $\{s \in S \mid b_s < v_s\}$, contradicting the maximality of U. ∎

(For an alternative characterization, see Welsh [1976].)

By (44.8) and (44.9) the rank of a is given by

(44.65) $r(a) = \min_{U \subseteq S}(a(S \setminus U) + f(U))$

(from this one may derive a 'submodular law' for r: $r(a \wedge b) + r(a \vee b) \leq r(a) + r(b)$, where $\wedge$ and $\vee$ are the meet and join in the lattice $(\mathbb{R}^S, \leq)$ (Edmonds [1970b])).

Since if P has integer vertices and a is integer, the intersection $P|a = \{x \in P \mid x \leq a\}$ is integer again, we know that for integer polymatroids (44.56) also holds if we restrict a and x to integer vectors. So if a is integer, then there exists an integer vector $x \leq a$ in P with $x(S) = r(a)$.

Theorem 44.5 yields an analogous characterization of extended polymatroids. Let f be a submodular set function on S with $f(\emptyset) = 0$. Choose $c \in \mathbb{R}^S_+$ such that

(44.66) $g(U) := f(U) + c(U)$

is nonnegative for all $U \subseteq S$. Clearly, g again is submodular, and $g(\emptyset) = 0$. Then the extended polymatroid EP_f associated with f and the polymatroid P_g associated with g are related by:

(44.67) $P_g = \{x \mid x \geq \mathbf{0}, x - c \in EP_f\} = (c + EP_f) \cap \mathbb{R}^S_+$.

Since P_g is a polymatroid, by (44.56) we know that EP_f satisfies:

(44.68) for each a in $\mathbb{R}^S$ there exists a number $r(a)$ such that each maximal vector x in $EP_f \cap \{x \in \mathbb{R}^S \mid x \leq a\}$ satisfies $x(S) = r(a)$.

One easily derives from Theorem 44.5 that (44.68) together with

(44.69) if $y \leq x \in EP_f$, then $y \in EP_f$,

characterizes the class of all extended polymatroids among the closed subsets of $\mathbb{R}^S$.

44.6e. Operations on submodular functions and polymatroids

The class of submodular set functions on a given set is closed under certain operations. Obviously, the sum of two submodular functions is submodular again. In particular, adding a constant t to all values of a submodular function maintains submodularity. Also the multiplication of a submodular function by a nonnegative scalar maintains submodularity. Moreover, if f is a nondecreasing submodular set function on S, and q is a real number, then the function f' given by $f'(U) := \min\{q, f(U)\}$ for $U \subseteq S$, is submodular again. (Monotonicity cannot be deleted, as is shown by taking $S := \{a, b\}$, $f(\emptyset) = f(S) = 1$, $f(\{a\}) = 0$, $f(\{b\}) = 2$, and $q = 1$.)

It follows that the class of all submodular set functions on S forms a convex cone C in $\mathbb{R}^{\mathcal{P}(S)}$. This cone is polyhedral as the constraints (44.1) form a finite set of linear inequalities defining C. Edmonds [1970b] raised the problem of determining the extreme rays of the cone of all nonnegative nondecreasing submodular set functions f on S with $f(\emptyset) = 0$. It is not difficult to show that the rank function r of a matroid M determines an extreme ray of this cone if and only if r is not the sum of the rank functions of two other matroids, i.e., if and only if M is the sum of a connected matroid and a number of loops. But these do not represent all extreme rays: if $S = \{1, \ldots, 5\}$ and $w(1) = 2, w(s) = 1$ for $s \in S \setminus \{1\}$, let $f(U) := \min\{3, w(U)\}$ for $U \subseteq S$; then f is on an extreme ray, but cannot be decomposed as the sum of rank functions of matroids (L. Lovász's example; cf. also Murty and Simon [1978] and Nguyen [1978]).

Lovász [1983c] observed that if f_1 and f_2 are submodular and $f_1 - f_2$ is nondecreasing, then $\min\{f_1, f_2\}$ is submodular.

Let f be a nonnegative submodular set function on S. Clearly, for any $\lambda \geq 0$ we have $P_{\lambda f} = \lambda P_f$ (where $\lambda P_f = \{\lambda x \mid x \in P_f\}$). If $q \geq 0$, and f' is given by $f'(U) = \min\{q, f(U)\}$ for $U \subseteq S$, then f' is submodular and

$$(44.70) \qquad P_{f'} = \{x \in P_f \mid x(S) \leq q\},$$

as can be checked easily. So the class of polymatroids is closed under intersections with affine halfspaces of the form $\{x \in \mathbb{R}^S \mid x(S) \leq q\}$, for $q \geq 0$.

Let f_1 and f_2 be nondecreasing submodular set functions on S, with $f_1(\emptyset) = f_2(\emptyset) = 0$, and associated polymatroids P_1 and P_2 respectively. Let P be the polymatroid associated with $f := f_1 + f_2$. Then (McDiarmid [1975c]):

Theorem 44.6. $P_{f_1+f_2} = P_{f_1} + P_{f_2}$.

Proof. It is easy to see that $P_{f_1+f_2} \supseteq P_{f_1} + P_{f_2}$. To prove the reverse inclusion, let x be a vertex of $P_{f_1+f_2}$. Then x has the form (44.49). Hence, by taking the same permutation π and the same $k, x = x_1 + x_2$ for certain vertices x_1 of P_{f_1} and x_2 of P_{f_2}. Since $P_{f_1} + P_{f_2}$ is convex it follows that $P_{f_1+f_2} = P_{f_1} + P_{f_2}$. ∎

In fact, if f_1 and f_2 are integer, each *integer* vector in $P_{f_1} + P_{f_2}$ is the sum of *integer* vectors in P_{f_1} and P_{f_2} — see Corollary 46.2c. Similarly, if f_1 and f_2 are integer, each integer vector in $EP_{f_1} + EP_{f_2}$ is the sum of integer vectors in EP_{f_1} and EP_{f_2}.

Faigle [1984a] derived from Theorem 44.6 that, for any submodular function f, if $x, y \in P_f$ and $x = x_1 + x_2$ with $x_1, x_2 \in P_f$, then there exist $y_1, y_2 \in P_f$ with

$y = y_1 + y_2$ and $x_1 + y_1, x_2 + y_2 \in P_f$. (Proof: $y \in P_f \subseteq P_{2f-x} = P_{f-x_1} + P_{f-x_2}$.) An integer version of this can be derived from Corollary 46.2c and generalizes (42.13).

If $M_1 = (S, \mathcal{I}_1)$ and $M_2 = (S, \mathcal{I}_2)$ are matroids, with rank functions r_1 and r_2 and corresponding independent set polytopes P_1 and P_2, respectively, then by Section 44.6c above, $P_1 + P_2$ is the convex hull of sums of incidence vectors of independent sets in M_1 and M_2. Hence the 0,1 vectors in $P_1 + P_2$ are just the incidence vectors of the sets $I_1 \cup I_2$, for $I_1 \in \mathcal{I}_1$ and $I_2 \in \mathcal{I}_2$. Therefore, the polyhedron

(44.71) $$(P_1 + P_2)|\mathbf{1} = \{x \in P_1 + P_2 \mid x \leq \mathbf{1}\}$$

is the convex hull of the independent sets of $M_1 \vee M_2$. By Theorem 44.6 and (44.45), it follows that the rank function r of $M_1 \vee M_2$ satisfies

(44.72) $$r(U) = \min_{T \subseteq U}(|U \setminus T| + r_1(T) + r_2(T))$$

for $U \subseteq S$. Thus we have derived the matroid union theorem (Corollary 42.1a).

44.6f. Duals of polymatroids

McDiarmid [1975c] described the following duality of polymatroids. Let P be the polymatroid associated with the nondecreasing submodular set function f on S with $f(\emptyset) = 0$ and let a be a vector in $\mathbb{R}^S$ with $a \geq x$ for all x in P (i.e., $a(s) \geq f(\{s\})$ for all s in S). Define

(44.73) $$f^*(U) := a(U) + f(S \setminus U) - f(S)$$

for $U \subseteq S$. One easily checks that f^* again is nondecreasing and submodular, and that $f^*(\emptyset) = 0$. We call f^* the *dual* of f (with respect to a). Then $f^{**} = f$ taking the second dual with respect to the same a, as follows immediately from (44.73).

Let P^* be the polymatroid associated with f^*, and call P^* the *dual* polymatroid of P (with respect to a). Now the maximal vertices of P and P^* are given by (44.49) by choosing $k = n$. It follows that x is a maximal vertex of P if and only if $a - x$ is a maximal vertex of P^*. Since the maximal vectors of a polymatroid form just the convex hull of the maximal vertices, we may replace in the previous sentence the word 'vertex' by 'vector'. So the set of maximal vectors of P^* arises from the set of maximal vectors of P by reflection in the point $\frac{1}{2}a$.

Clearly, duals of matroids correspond in the obvious way to duals of the related polymatroids (with respect to the vector $\mathbf{1}$).

44.6g. Induction of polymatroids

Let $G = (V, E)$ be a bipartite graph, with colour classes S and T. Let f be a nondecreasing submodular set function on S with $f(\emptyset) = 0$, and define

(44.74) $$g(U) := f(N(U))$$

for $U \subseteq T$ (cf. Section 44.1a). (As usual, $N(U)$ denotes the set of vertices not in U adjacent to at least one vertex in U.)

The function g again is nondecreasing and submodular. Similarly to Rado's theorem (Corollary 41.1c), one may prove that a vector x belongs to P_g if and only if there exist $y \in \mathbb{R}^E_+$ and $z \in P_f$ such that

$$
\begin{aligned}
(44.75) \qquad & y(\delta(t)) = x_t \quad (t \in T), \\
& y(\delta(s)) = z_s \quad (s \in S).
\end{aligned}
$$

Moreover, if f and g are integer, we can take y and z to be integer. This procedure gives an 'induction' of polymatroids through bipartite graphs, and yields 'Rado's theorem for polymatroids' (cf. McDiarmid [1975c]).

In case f is the rank function of a matroid on S, a 0,1 vector x belongs to P_g if and only if there exists a matching in G whose end vertices in S form an independent set of the matroid, and the end vertices in T have x as incidence vector. So these 0,1 vectors determine a matroid on T, with rank function r given by

$$(44.76) \qquad r(U) = \min_{W \subseteq U} (|U \setminus W| + f(N(W)))$$

for $U \subseteq T$ (cf. (44.45) and (44.74)).

Another extension is the following. Let $D = (V, A)$ be a directed graph and let V be partitioned into classes S and T. Let furthermore a 'capacity' function $c : A \to \mathbb{R}_+$ be given. Define the set function g on T by

$$(44.77) \qquad g(U) := c(\delta^{\text{out}}(U))$$

for $U \subseteq T$, where $\delta^{\text{out}}(U)$ denotes the set of arcs leaving U. Then g is nonnegative and submodular, and it may be derived straightforwardly from the max-flow min-cut theorem (Theorem 10.3) that a vector x in $\mathbb{R}_+^T$ belongs to P_g if and only if there exist $T - S$ paths $Q_1, \ldots, Q_k$ and nonnegative numbers $\lambda_1, \ldots, \lambda_k$ (for some k), such that

$$(44.78) \qquad \sum_{i=1}^{k} \lambda_i \chi^{AQ_i} \leq c \text{ and } \sum_{i=1}^{k} \lambda_i \chi^{b(Q_i)} = x,$$

where $b(Q_i)$ is the beginning vertex of Q_i. If the c and x are integer, we can take also the λ_i integer.

Here the function g in general is not nondecreasing, but the value

$$(44.79) \qquad \bar{g}(U) = \min\{g(W) \mid U \subseteq W \subseteq T\}$$

of the associated nondecreasing submodular function (cf. (44.35)) is equal to the minimum capacity of a cut separating U and S, which is equal to the maximum amount of flow from U to S, subject to the capacity function c (by the max-flow min-cut theorem).

In an analogous way, one can construct polymatroids by taking vertex-capacities instead of arc-capacities. Moreover, the notion of induction of polymatroids through bipartite graphs can be extended in a natural way to the induction of polymatroids through directed graphs (cf. McDiarmid [1975c], Schrijver [1978]).

44.6h. Lovász's generalization of Kőnig's matching theorem

Lovász [1970a] gave the following generalization of Kőnig's matching theorem (Theorem 16.2).

For a graph $G = (V, E)$, $U \subseteq V$, and $F \subseteq E$, let $N_F(U)$ denote the set of vertices not in U that are adjacent in (V, F) to at least one vertex in U. Kőnig's matching theorem follows by taking $g(X) := |X|$ in the following theorem.

Theorem 44.7. *Let $G = (V, E)$ be a simple bipartite graph, with colour classes S and T. Let g be a supermodular set function on S, such that $g(\{v\}) \geq 0$ for each $v \in S$ and such that*

(44.80) $\quad g(U \cup \{v\}) \leq g(U) + g(\{v\})$ *for nonempty* $U \subseteq S$ *and* $v \in S \setminus U$.

Then E has a subset F with $\deg_F(v) = g(\{v\})$ for each $v \in V$ and $|N_F(U)| \geq g(U)$ for each nonempty $U \subseteq S$ if and only if $|N_E(U)| \geq g(U)$ for each nonempty $U \subseteq S$.

Proof. Necessity being trivial, we show sufficiency. Choose $F \subseteq E$ such that

(44.81) $\quad |N_F(U)| \geq g(U)$

for each nonempty $U \subseteq S$, with $|F|$ as small as possible. We show that F is as required.

Suppose to the contrary that $\deg_F(v) > g(\{v\})$ for some $v \in S$. By the minimality of F, for each edge $e = vw \in F$, there is a subset U_e of S with $v \in U_e$, $|N_F(U_e)| = g(U_e)$, and $w \notin N_F(U_e \setminus \{v\})$. Since the function $|N_F(U)|$ is submodular, the intersection U of the U_e over $e \in \delta(v)$ satisfies $|N_F(U)| = g(U)$ (using (44.81)). Then no neighbour w of v is adjacent to U. Hence $N_F(v)$ and $N_F(U \setminus \{v\})$ are disjoint. Moreover, $U \neq \{v\}$, since $N_F(U) = g(U)$ and $N_F(\{v\}) > g(v)$. This gives the contradiction

(44.82) $\quad g(U) \leq g(U \setminus \{v\}) + g(\{v\}) < |N_F(U \setminus \{v\})| + |N_F(v)| = |N_F(U)|.$ ∎

For a derivation of this theorem with the Edmonds-Giles method, see Frank and Tardos [1989].

44.6i. Further notes

Edmonds [1970b] and D.A. Higgs (as mentioned in Edmonds [1970b]) observed that if f is a set function on a set S, we can define recursively a submodular function $\bar{f}$ as follows:

(44.83) $\quad \bar{f}(T) := \min\{f(T), \min(\bar{f}(S_1) + \bar{f}(S_2) - \bar{f}(S_1 \cap S_2))\},$

where the second minimum ranges over all pairs S_1, S_2 of proper subsets of T with $S_1 \cup S_2 = T$.

Lovász [1983c] gave the following characterization of submodularity in terms of convexity. Let f be a set function on S and define for each $c \in \mathbb{R}^S_+$

(44.84) $$\hat{f}(c) := \sum_{i=1}^{k} \lambda_i f(U_i),$$

where $\emptyset \neq U_1 \subset U_2 \subset \cdots \subset U_k \subseteq S$ and $\lambda_1, \ldots, \lambda_k > 0$ are such that $c = \sum_{i=1}^{k} \lambda_i \chi^{U_i}$. Then f is submodular if and only if $\hat{f}$ is convex. Similarly, f is supermodular if and only if $\hat{f}$ is concave. Related is the 'subdifferential' of a submodular function, investigated by Fujishige [1984d].

Korte and Lovász [1985c] and Nakamura [1988a] studied polyhedral structures where the greedy algorithm applies. Federgruen and Groenevelt [1986] extended the greedy method for polymatroids to 'weakly concave' objective functions (instead of linear functions). (Related work was reported by Bhattacharya, Georgiadis, and

Tsoucas [1992].) Nakamura [1993] extended polymatroids and submodular functions to Δ-polymatroids and Δ-submodular functions.

Gröflin and Liebling [1981] studied the following example of 'transversal polymatroids'. Let $G = (V, E)$ be an undirected graph, and define the submodular set function f on E by $f(F) := \left|\bigcup F\right|$ for $F \subseteq E$. Then the vertices of the associated polymatroid are all $\{0, 1, 2\}$ vectors x in $\mathbb{R}^E$ with the property that the set $F := \{e \in E \mid x_e \geq 1\}$ forms a forest each component of which contains at most one edge e with $x_e = 2$. If x is a maximal vertex, then each component contains exactly one edge e with $x_e = 2$.

Narayanan [1991] studied, for a given submodular function f on S, the lattice of all partitions $\mathcal{P}$ of S into nonempty sets such that there exists a $\lambda \in \mathbb{R}$ for which $\mathcal{P}$ attains $\min \sum_{U \in \mathcal{P}}(f(U) - \lambda)$ (taken over all partitions $\mathcal{P}$). Fujishige [1980b] studied minimum values of submodular functions.

For results on the (NP-hard) problems of *maximizing* a submodular function and of submodular set cover, see Fisher, Nemhauser, and Wolsey [1978], Nemhauser and Wolsey [1978,1981], Nemhauser, Wolsey, and Fisher [1978], Wolsey [1982a,1982b], Conforti and Cornuéjols [1984], and Fujito [1999].

Cunningham [1983], Fujishige [1983], and Nakamura [1988c] presented decomposition theories for submodular functions. Benczúr and Frank [1999] considered covering symmetric supermodular functions by graphs.

For surveys and books on polymatroids and submodular functions, see McDiarmid [1975c], Welsh [1976], Lovász [1983c], Lawler [1985], Nemhauser and Wolsey [1988], Fujishige [1991], Narayanan [1997], and Murota [2002]. For a survey on applications of submodular functions, see Frank [1993a].

Historically, submodular functions arose in lattice theory (Bergmann [1929], Birkhoff [1933]), while submodularity of the rank function of a matroid was shown by Bergmann [1929] and Whitney [1935]. Choquet [1951,1955] and Kelley [1959] studied submodular functions in relation to the Newton capacity and to measures in Boolean algebras. The relevance of submodularity for optimization was revealed by Edmonds [1970b].

Several alternative names have been proposed for submodular functions, like sub-valuation (Choquet [1955]), β-function (Edmonds [1970b]), and ground set rank function (McDiarmid [1975c]). The set of integer vectors in an integer polymatroid was called a hypermatroid by Helgason [1974] and Lovász [1977c]. A generalization of polymatroids (called supermatroids) was studied by Dunstan, Ingleton, and Welsh [1972].

Chapter 45

Submodular function minimization

This chapter describes a strongly polynomial-time algorithm to find the minimum value of a submodular function. It suffices that the submodular function is given by a value giving oracle.
One application of submodular function minimization is optimizing over the intersection of two polymatroids. This will be discussed in Chapter 47.

45.1. Submodular function minimization

It was shown by Grötschel, Lovász, and Schrijver [1981] that the minimum value of a rational-valued submodular set function f on S can be found in polynomial time, if f is given by a value giving oracle and an upper bound B is given on the numerators and denominators of the values of f. The running time is bounded by a polynomial in $|S|$ and $\log B$. This algorithm is based on the ellipsoid method: we can assume that $f(\emptyset) = 0$ (by resetting $f(U) := f(U) - f(\emptyset)$ for all $U \subseteq S$); then with the greedy algorithm, we can optimize over EP_f in polynomial time (Corollary 44.3b), hence the separation problem for EP_f is solvable in polynomial time, hence also the separation problem for

$$(45.1) \qquad P := EP_f \cap \{x \mid x \leq \mathbf{0}\},$$

and therefore also the optimization problem for P. Now the maximum value of $x(S)$ over P is equal to the minimum value of f (by (44.8), (44.9), and (44.34)).

Having a polynomial-time method to find the minimum value of a submodular function, we can turn it into a polynomial-time method to find a subset T of S minimizing $f(T)$: For each $s \in S$, we can determine if the minimum value of f over all subsets of S is equal to the minimum value of f over subsets of $S \setminus \{s\}$. If so, we reset $S := S \setminus \{s\}$. Doing this for all elements of S, we are left with a set T minimizing f over all subsets of (the original) S.

Grötschel, Lovász, and Schrijver [1988] showed that this algorithm can be turned into a strongly polynomial-time method. Cunningham [1985b] gave a combinatorial, pseudo-polynomial-time algorithm for minimizing a submodular function f (polynomial in the size of the underlying set and the maximum absolute value of f (assuming f to be integer)). Inspired by Cunningham's method, combinatorial strongly polynomial-time algorithms were found by Iwata, Fleischer, and Fujishige [2000,2001] and Schrijver [2000a]. We will describe the latter algorithm.

45.2. Orders and base vectors

Let f be a submodular set function on a set S. In finding the minimum value of f, we can assume $f(\emptyset) = 0$, as resetting $f(U) := f(U) - f(\emptyset)$ for all $U \subseteq S$ does not change the problem. So throughout we assume that $f(\emptyset) = 0$.

Moreover, we assume that f is given by a *value giving oracle*, that is, an oracle that returns $f(U)$ for any given subset U of S. We also assume that the numbers returned by the oracle are rational (or belong to any ordered field in which we can perform the elementary arithmetic operations algorithmically).

Recall that the *base polytope* B_f of f is defined as the set of base vectors of f:

(45.2) $$B_f := \{x \in \mathbb{R}^S \mid x(U) \leq f(U) \text{ for all } U \subseteq S, x(S) = f(S)\}.$$

Consider any total order $\prec$ on S.[35] For any $v \in S$, denote

(45.3) $$v_\prec := \{u \in S \mid u \prec v\}.$$

Define a vector $b^\prec$ in $\mathbb{R}^S$ by:

(45.4) $$b^\prec(v) := f(v_\prec \cup \{v\}) - f(v_\prec)$$

for $v \in S$. Theorem 44.3 implies that $b^\prec$ belongs to B_f.

Note that $b^\prec(U) = f(U)$ for each lower ideal U of $\prec$ (where a *lower ideal* of $\prec$ is a subset U of S such that if $v \in U$ and $u \prec v$, then $u \in U$).

45.3. A subroutine

In this section we describe a subroutine that is important in the algorithm. It replaces a total order $\prec$ by other total orders, thereby reducing some interval $(s,t]_\prec$, where

(45.5) $$(s,t]_\prec := \{v \mid s \prec v \preceq t\}$$

for $s, t \in S$.

Let $\prec$ be a total order on S. For any $s, u \in S$ with $s \prec u$, let $\prec^{s,u}$ be the total order on S obtained from $\prec$ by resetting $v \prec u$ to $u \prec v$ for each

[35] As usual, we use $\prec$ for strict inequality and $\preceq$ for nonstrict inequality. We refer to the order by the strict inequality sign $\prec$.

v satisfying $s \preceq v \prec u$. Thus in the ordering, we move u to the position just before s. Hence $(s,t]_{\prec^{s,u}} = (s,t]_{\prec} \setminus \{u\}$ if $u \in (s,t]_{\prec}$.

We show that there is a strongly polynomial-time subroutine that

(45.6) for any $s, t \in S$ with $s \prec t$, finds a $\delta \geq 0$ and describes $b^{\prec} + \delta(\chi^t - \chi^s)$ as a convex combination of the $b^{\prec^{s,u}}$ for $u \in (s,t]_{\prec}$.

To describe the subroutine, we can assume that $b^{\prec} = \mathbf{0}$, by replacing (temporarily) $f(U)$ by $f(U) - b^{\prec}(U)$ for each $U \subseteq S$.

We investigate the signs of the vector $b^{\prec^{s,u}}$. We show that for each $v \in S$:

(45.7) $b^{\prec^{s,u}}(v) \leq 0$ if $s \preceq v \prec u$,
$b^{\prec^{s,u}}(v) \geq 0$ if $v = u$,
$b^{\prec^{s,u}}(v) = 0$ otherwise.

To prove this, observe that if $T \subseteq U \subseteq S$, then for any $v \in S \setminus U$ we have by the submodularity of f:

(45.8) $f(U \cup \{v\}) - f(U) \leq f(T \cup \{v\}) - f(T)$.

To see (45.7), if $s \preceq v \prec u$, then by (45.8),

(45.9) $$\begin{aligned} b^{\prec^{s,u}}(v) &= f(v_{\prec^{s,u}} \cup \{v\}) - f(v_{\prec^{s,u}}) \leq f(v_{\prec} \cup \{v\}) - f(v_{\prec}) \\ &= b^{\prec}(v) = 0, \end{aligned}$$

since $v_{\prec^{s,u}} = v_{\prec} \cup \{u\} \supset v_{\prec}$.

Similarly,

(45.10) $$\begin{aligned} b^{\prec^{s,u}}(u) &= f(u_{\prec^{s,u}} \cup \{u\}) - f(u_{\prec^{s,u}}) \geq f(u_{\prec} \cup \{u\}) - f(u_{\prec}) \\ &= b^{\prec}(u) = 0, \end{aligned}$$

since $u_{\prec^{s,u}} = s_{\prec} \subset u_{\prec}$.

Finally, if $v \prec s$ or $u \prec v$, then $v_{\prec^{s,u}} = v_{\prec}$, and hence $b^{\prec^{s,u}}(v) = b^{\prec}(v) = 0$. This shows (45.7).

By (45.7), the matrix $M = (b^{\prec^{s,u}}(v))_{u,v}$ with rows indexed by $u \in (s,t]_{\prec}$ and columns indexed by $v \in S$, in the order given by $\prec$, has the following, partially triangular, shape, where $+$ means that the entry is ≥ 0, and $-$ that the entry is ≤ 0:

$$\begin{array}{c|cccccccccccccc}
 & & & & s & & & & & & & t & & & \\
\hline
 & 0 & \cdots & 0 & - & + & 0 & \cdots & \cdots & \cdots & 0 & 0 & 0 & \cdots & 0 \\
 & \vdots & & \vdots & - & - & + & \ddots & & & \vdots & \vdots & \vdots & & \vdots \\
 & \vdots & & \vdots & - & - & - & \ddots & \ddots & & \vdots & \vdots & \vdots & & \vdots \\
 & \vdots & & \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \vdots & & \vdots \\
 & \vdots & & \vdots & \vdots & \vdots & \vdots & & \ddots & \ddots & 0 & 0 & \vdots & & \vdots \\
 & \vdots & & \vdots & \vdots & \vdots & \vdots & & & \ddots & + & 0 & \vdots & & \vdots \\
t & 0 & \cdots & 0 & - & - & - & \cdots & \cdots & \cdots & - & + & 0 & \cdots & 0
\end{array}$$

As each row of M represents a vector $b^{\prec^{s,u}}$, to obtain (45.6) we must describe $\delta(\chi^t - \chi^s)$ as a convex combination of the rows of M, for some $\delta \geq 0$.

We call the + entries in the matrix the 'diagonal' elements. Now for each row of M, the sum of its entries is 0, as $b^{\prec^{s,u}}(S) = f(S) = b^{\prec}(S) = 0$. Hence, if a 'diagonal' element $b^{\prec^{s,u}}(u)$ is equal to 0 for some $u \in (s,t]_{\prec}$, then the corresponding row of M is all-zero. So in this case we can take $\delta = 0$ in (45.6).

If $b^{\prec^{s,u}}(u) > 0$ for each $u \in (s,t]_{\prec}$ (that is, if each 'diagonal' element is strictly positive), then $\chi^t - \chi^s$ can be described as a nonnegative combination of the rows of M (by the sign pattern of M and since the entries in each row of M add up to 0). Hence $\delta(\chi^t - \chi^s)$ is a convex combination of the rows of M for some $\delta > 0$, yielding again (45.6).

45.4. Minimizing a submodular function

We now describe the algorithm to find the minimum value of a submodular set function f on S. We assume $f(\emptyset) = 0$ and $S = \{1, \ldots, n\}$.

We iteratively update a vector $x \in B_f$, given as a convex combination

(45.11) $\quad x = \lambda_1 b^{\prec_1} + \cdots + \lambda_k b^{\prec_k},$

where the $\prec_i$ are total orders of S, and where the λ_i are positive and sum to 1. Initially, we choose an arbitrary total order $\prec$ and set $x := b^{\prec}$ (so $k = 1$ and $\prec_1 = \prec$).

We describe the iteration. Consider the directed graph $D = (S, A)$, with

(45.12) $\quad A := \{(u,v) \mid \exists i = 1, \ldots, k : u \prec_i v\}.$

Define

(45.13) $\quad P := \{v \in S \mid x(v) > 0\}$ and $N := \{v \in S \mid x(v) < 0\}.$

Case 1: *D* has no path from *P* to *N*. Then let U be the set of vertices of D that can reach N by a directed path. So $N \subseteq U$ and $U \cap P = \emptyset$; that is, U contains all negative components of x and no positive components. Hence $x(W) \geq x(U)$ for each $W \subseteq S$. As no arcs of D enter U, U is a lower ideal of $\prec_i$, and hence $b^{\prec_i}(U) = f(U)$, for each $i = 1, \ldots, k$. Therefore, for each $W \subseteq S$:

(45.14) $$f(U) = \sum_{i=1}^{k} \lambda_i b^{\prec_i}(U) = x(U) \leq x(W) \leq f(W).$$

So U minimizes f.

Case 2: *D* has a path from *P* to *N*. Let $d(v)$ denote the distance in D from P to v (= minimum number of arcs in a directed path from P to v). Set $d(v) := \infty$ if v is not reachable from P. Choose $s, t \in S$ as follows.

Let t be the element in N reachable from P with $d(t)$ maximum, such that t is largest. Let s be the element with $(s,t) \in A$, $d(s) = d(t) - 1$, and s largest. Let α be the maximum of $|(s,t]_{\prec_i}|$ over $i = 1, \ldots, k$. Reorder indices such that $|(s,t]_{\prec_1}| = \alpha$.

By (45.6), we can find $\delta \geq 0$ and describe

(45.15) $$b^{\prec_1} + \delta(\chi^t - \chi^s)$$

as a convex combination of the $b^{\prec_1^{s,u}}$ for $u \in (s,t]_{\prec_1}$. Then with (45.11) we obtain

(45.16) $$y := x + \lambda_1 \delta(\chi^t - \chi^s)$$

as a convex combination of $b^{\prec_i}$ $(i = 2, \ldots, k)$ and $b^{\prec_1^{s,u}}$ $(u \in (s,t]_{\prec_1})$.

Let x' be the point on the line segment $\overline{xy}$ closest to y satisfying $x'(t) \leq 0$. (So $x'(t) = 0$ or $x' = y$.) We can describe x' as a convex combination of $b^{\prec_i}$ $(i = 1, \ldots, k)$ and $b^{\prec_1^{s,u}}$ $(u \in (s,t]_{\prec_1})$. Moreover, if $x'(t) < 0$, then we can do without $b^{\prec_1}$.

We reduce the number of terms in the convex decomposition of x' to at most $|S|$ by linear algebra: any affine dependence of the vectors in the decomposition yields a reduction of the number of terms in the decomposition, as in the standard proof of Carathéodory's theorem (subtract an appropriate multiple of the linear expression giving the affine dependence, from the linear expression giving the convex combination, such that all coefficients remain nonnegative, and at least one becomes 0). As all $b^{\prec}$ belong to a hyperplane, this reduces the number of terms to at most $|S|$.

Then reset $x := x'$ and iterate. This finishes the description of the algorithm.

45.5. Running time of the algorithm

We show that the number of iterations is at most $|S|^6$. Consider any iteration. Let

(45.17) $$\beta := \text{number of } i \in \{1, \ldots, k\} \text{ with } |(s,t]_{\prec_i}| = \alpha.$$

Let x', d', A', P', N', t', s', α', β' be the objects x, d, A, P, N, t, s, α, β in the next iteration (if any). Then

(45.18) for all $v \in S$, $d'(v) \geq d(v)$,

and

(45.19) if $d'(v) = d(v)$ for all $v \in S$, then $(d'(t'), t', s', \alpha', \beta')$ is lexicographically less than $(d(t), t, s, \alpha, \beta)$.

Since each of $d(t), t, s, \alpha, \beta$ is at most $|S|$, and since (if $d(v)$ is unchanged for all v) there are at most $|S|$ pairs $(d(t), t)$, (45.19) implies that in at most $|S|^4$

iterations $d(v)$ increases for at least one v. Any fixed v can have at most $|S|$ such increases, and hence the number of iterations is at most $|S|^6$.

In order to prove (45.18) and (45.19), notice that

(45.20) for each arc $(v, w) \in A' \setminus A$ we have $s \preceq_1 w \prec_1 v \preceq_1 t$.

Indeed, as $(v, w) \notin A$ we have $w \prec_1 v$. As $(v, w) \in A'$, we have $v \prec_1^{s,u} w$ for some $u \in (s, t]_{\prec_1}$. Hence the definition of $\prec_1^{s,u}$ gives $v = u$ and $s \preceq_1 w \prec_1 u$. This shows (45.20).

If (45.18) does not hold, then $A' \setminus A$ contains an arc (v, w) with $d(w) \geq d(v) + 2$ (using that $P' \subseteq P$). By (45.20), $s \preceq_1 w \prec_1 v \preceq_1 t$, and so $d(w) \leq d(s) + 1 = d(t) \leq d(v) + 1$, a contradiction. This shows (45.18).

To prove (45.19), assume that $d'(v) = d(v)$ for all $v \in S$. As $x'(t') < 0$, we have $x(t') < 0$ or $t' = s$. So by our criterion for choosing t (maximizing $(d(t), t)$ lexicographically), and since $d(s) < d(t)$, we know that $d(t') \leq d(t)$, and that if $d(t') = d(t)$, then $t' \leq t$.

Next assume that moreover $d(t') = d(t)$ and $t' = t$. As $(s', t) \in A'$, and as (by (45.20)) $A' \setminus A$ contains no arc entering t, we have $(s', t) \in A$, and so $s' \leq s$, by the maximality of s.

Finally assume that moreover $s' = s$. As $(s, t]_{\prec_1^{s,u}}$ is a proper subset of $(s, t]_{\prec_1}$ for each $u \in (s, t]_{\prec_1}$, we know that $\alpha' \leq \alpha$. Moreover, if $\alpha' = \alpha$, then $\beta' < \beta$, since $\prec_1$ does not occur anymore among the linear orders making the convex combination, as $x'(t) < 0$. This proves (45.19).

We therefore have proved:

Theorem 45.1. *Given a submodular function f by a value giving oracle, a set U minimizing $f(U)$ can be found in strongly polynomial time.*

Proof. See above. ∎

This algorithm performs the elementary arithmetic operations on function values, including multiplication and division (in order to solve certain systems of linear equations). One would wish to have a 'fully combinatorial' algorithm, in which the function values are only compared, added, and subtracted. The existence of such an algorithm was shown by Iwata [2002a, 2002c], by extending the algorithm of Iwata, Fleischer, and Fujishige [2000, 2001].

Notes. In the algorithm, we have chosen t and s largest possible, in some fixed order of S. To obtain the above running time bound it only suffices to choose t and s in a consistent way. That is, if the set of choices for t is the same as in the previous iteration, then we should choose the same t — and similarly for s. This roots in the idea of 'consistent breadth-first search' of Schönsleben [1980].

The observation that the number of iterations in the algorithm of Section 45.4 is $O(|S|^6)$ instead of $O(|S|^7)$ is due to L.K. Fleischer. Vygen [2002] showed that the number of iterations can in fact be bounded by $O(|S|^5)$.

The algorithm described above has been speeded up by Fleischer and Iwata [2000,2002], by incorporating a push-relabel type of iteration based on approximate distances instead of precise distances (like Goldberg's method for maximum flow, given in Section 10.7). Iwata [2002b] combined the approaches of Iwata, Fleischer, and Fujishige [2000,2001] and Schrijver [2000a] to obtain a faster algorithm. A descent method for submodular function minimization based on an oracle for membership of the base polytope was given by Fujishige and Iwata [2002].

Surveys and background on submodular function minimization are given by Fleischer [2000b] and McCormick [2001].

45.6. Minimizing a symmetric submodular function

A set function f on S is called *symmetric* if $f(U) = f(S \setminus U)$ for each $U \subseteq S$. The minimum of a symmetric submodular function f is attained by $\emptyset$, since for each $U \subseteq S$ one has

(45.21) $$2f(U) = f(U) + f(S \setminus U) \geq f(\emptyset) + f(S) = 2f(\emptyset).$$

By extending a method of Nagamochi and Ibaraki [1992b] for finding the minimum nonempty cut in an undirected graph, Queyranne [1995,1998] gave an easy combinatorial algorithm to find a nonempty proper subset U of S minimizing $f(U)$, where f is given by a value giving oracle. We may assume that $f(\emptyset) = f(S) = 0$, by resetting $f(U) := f(U) - f(\emptyset)$ for all $U \subseteq S$.

Call an ordering $s_1, \ldots, s_n$ of the elements of S a *legal order* of S for f, if, for each $i = 1, \ldots, n$,

(45.22) $$f(\{s_1, \ldots, s_{i-1}, x\}) - f(\{x\})$$

is minimized over $x \in S \setminus \{s_1, \ldots, s_{i-1}\}$ by $x = s_i$. One easily finds a legal order, by $O(|S|^2)$ oracle calls (for the value of f).

Now the algorithm is (where a set U *splits* a set X if both $X \cap U$ and $X \setminus U$ are nonempty):

(45.23) Find a legal order $(s_1, \ldots, s_n)$ of S for f.
Determine (recursively) a nonempty proper subset T of S not splitting $\{s_{n-1}, s_n\}$, minimizing $f(T)$. (This can be done by identifying s_{n-1} and s_n.)
Then the minimum value of $f(U)$ over nonempty proper subsets U of S is equal to $\min\{f(T), f(\{s_n\})\}$.

The correctness of the algorithm follows from, for $n \geq 2$:

(45.24) $$f(U) \geq f(\{s_n\}) \text{ for each } U \subseteq S \text{ splitting } \{s_{n-1}, s_n\}.$$

This can be proved as follows. Define $t_0 := s_1$. For $i = 1, \ldots, n-1$, define $t_i := s_j$, where j is the smallest index such that $j > i$ and such that U splits $\{s_i, s_j\}$. For $i = 0, \ldots, n$, let $U_i := \{s_1, \ldots, s_i\}$. Note that for each $i = 1, \ldots, n-1$ one has

(45.25) $$f(U_{i-1} \cup \{t_i\}) - f(\{t_i\}) \geq f(U_{i-1} \cup \{t_{i-1}\}) - f(\{t_{i-1}\}),$$

since if $t_{i-1} = t_i$ this is trivial, and if $t_{i-1} \neq t_i$, then $t_{i-1} = s_i$, in which case (45.25) follows from the legality of the order.

Moreover, for each $i = 1, \ldots, n-1$ (setting $\overline{U} := S \setminus U$):

(45.26) $$\begin{aligned} & f(U_i \cup U) - f(U_{i-1} \cup U) + f(U_i \cup \overline{U}) - f(U_{i-1} \cup \overline{U}) \\ & \leq f(U_i \cup \{t_i\}) - f(U_{i-1} \cup \{t_i\}). \end{aligned}$$

In proving this, we may assume (by symmetry of U and $\overline{U}$) that $s_i \in \overline{U}$. Then $U_i \cup \overline{U} = U_{i-1} \cup \overline{U}$ and $t_i \in U$. So $f(U_i \cup \{t_i\}) + f(U_{i-1} \cup U) \geq f(U_{i-1} \cup \{t_i\}) + f(U_i \cup U)$, by submodularity. This gives (45.26).

Then we have:

(45.27) $$\begin{aligned} & f(s_n) - 2f(U) \\ & = f(U_{n-1} \cup U) + f(U_{n-1} \cup \overline{U}) - f(U_0 \cup U) - f(U_0 \cup \overline{U}) \\ & = \sum_{i=1}^{n-1} (f(U_i \cup U) - f(U_{i-1} \cup U) + f(U_i \cup \overline{U}) - f(U_{i-1} \cup \overline{U})) \\ & \leq \sum_{i=1}^{n-1} (f(U_i \cup \{t_i\}) - f(U_{i-1} \cup \{t_i\})) \\ & \leq \sum_{i=1}^{n-1} (f(U_i \cup \{t_i\}) - f(U_{i-1} \cup \{t_{i-1}\}) + f(\{t_{i-1}\}) - f(\{t_i\})) \\ & = f(U_{n-1} \cup \{t_{n-1}\}) - f(\{t_{n-1}\}) - f(\{t_0\}) + f(\{t_0\}) = -f(s_n) \end{aligned}$$

(where the first inequality follows from (45.26), and the second inequality from (45.25)). This shows (45.24).

Notes. Fujishige [1998] gave an alternative correctness proof. Nagamochi and Ibaraki [1998] extended the algorithm to minimizing submodular functions f satisfying

(45.28) $$f(T) + f(U) \geq f(T \setminus U) + f(U \setminus T)$$

for all $T, U \subseteq S$. Rizzi [2000b] gave an extension.

45.7. Minimizing a submodular function over the odd sets

From the strong polynomial-time solvability of submodular function minimization, one can derive that also a set of odd cardinality minimizing f (over the odd sets) is solvable in strongly polynomial time (Grötschel, Lovász, and Schrijver [1981,1984a,1988] (the second paper corrects a wrong argument given in the first paper)).

Theorem 45.2. *Given a submodular set function f on S (by a value giving oracle) and a nonempty subset T of S, one can find in strongly polynomial time a set $W \subseteq S$ minimizing $f(W)$ over W with $|W \cap T|$ odd.*

Proof. The case T odd can be reduced to the case T even as follows. Find for each $t \in T$ a subset W_t of $S - t$ with $W_t \cap (T - t)$ odd, and minimizing $f(W_t)$. Moreover, find a subset U of S minimizing $f(U)$ over $U \supseteq T$. Then a set that attains the minimum among $f(U)$ and the $f(W_t)$, is an output as required.

So we can assume that T is even. We describe a recursive algorithm. Say that a set U *splits* T if both $T \cap U$ and $T \setminus U$ are nonempty. First find a set U minimizing $f(U)$ over all subsets U of S splitting T. This can be done by finding for all $s, t \in T$ a set $U_{s,t}$ minimizing $f(U_{s,t})$ over all subsets of S containing s and not containing t (this amounts to submodular function minimization), and taking for U a set that minimizes $f(U_{s,t})$ over all such s, t.

If $U \cap T$ is odd, we output $W := U$. If $U \cap T$ is even, then recursively we find a set X minimizing $f(X)$ over all X with $X \cap (T \cap U)$ odd, and not splitting $T \setminus U$. This can be done by shrinking $T \setminus U$ to one element. Also, recursively we can find a set Y minimizing $f(Y)$ over all Y with $Y \cap (T \setminus U)$ odd, and not splitting $T \cap U$. Output an X or Y attaining the minimum of $f(X)$ and $f(Y)$.

This gives a strongly polynomial-time algorithm as the total number of recursive calls is at most $|T| - 2$ (since $2 + (|T \cap U| - 2) + (|T \setminus U| - 2) = |T| - 2$).

To see the correctness, let W minimize $f(W)$ over those W with $|W \cap T|$ odd. Suppose that $f(W) < f(X)$ and $f(W) < f(Y)$. As $f(W) < f(X)$, W splits $T \setminus U$, and hence $W \cup U$ splits T. Similarly, $f(W) < f(Y)$ implies that $W \cap U$ splits T.

Since $W \cap T$ is odd and $U \cap T$ is even, either $(W \cap U) \cap T$ or $(W \cup U) \cap T$ is odd.

If $(W \cap U) \cap T$ is odd, then $f(W \cap U) \geq f(W)$ (as W minimizes $f(W)$ over W with $W \cap T$ odd) and $f(W \cup U) \geq f(U)$ (as $W \cup U$ splits T and as U minimizes $f(U)$ over U splitting T). Hence, by the submodularity of f, $f(W \cap U) = f(W)$. Since $(W \cap U) \cap (T \cap U) = (W \cap U) \cap T$ is odd and since $W \cap U$ does not split $T \setminus U$, we have $f(W) = f(W \cap U) \geq f(X)$, contradicting our assumption.

If $(W \cup U) \cap T$ is odd, a symmetric argument gives a contradiction. ∎

This generalizes the strong polynomial-time solvability of finding a minimum-capacity odd cut in a graph, proved by Padberg and Rao [1982] (Corollary 25.6a). For a further generalization, see Section 49.11a.

Chapter 46

Polymatroid intersection

The intersection of two polymatroids behaves as nice as the intersection of two matroids, as was shown by Edmonds again. We study in this chapter min-max relations, polyhedral characterizations, and total dual integrality results. In the next chapter we go over to the algorithmic questions.

46.1. Box-total dual integrality of polymatroid intersection

We saw in Section 44.2 that the greedy algorithm yields a proof that an integer-valued submodular function gives an integer polymatroid. The interest of polymatroids for combinatorial optimization is enlarged by the fundamental result of Edmonds [1970b] that also the intersection of two integer polymatroids is integer, thus generalizing the matroid intersection theorem. In order to obtain this result, we first show a more general theorem (also due to Edmonds [1970b]).

Consider, for submodular set functions f_1, f_2 on S, the system:

(46.1) $$\begin{aligned} x(U) &\leq f_1(U) \quad \text{for } U \subseteq S, \\ x(U) &\leq f_2(U) \quad \text{for } U \subseteq S. \end{aligned}$$

Then:

Theorem 46.1. *If f_1 and f_2 are submodular, then* (46.1) *is box-TDI.*

Proof. Choose $w : S \to \mathbb{R}$. Let y_1, y_2 attain

(46.2) $$\begin{aligned} \min\{\sum_{U \subseteq S} &(y_1(U)f_1(U) + y_2(U)f_2(U)) \mid \\ &y_1, y_2 \in \mathbb{R}_+^{\mathcal{P}(S)}, \sum_{U \subseteq S}(y_1(U) + y_2(U))\chi^U = w\}. \end{aligned}$$

For $i = 1, 2$, define $w_i : S \to \mathbb{R}$ by

(46.3) $$w_i := \sum_{U \subseteq S} y_i(U)\chi^U.$$

Then y_i attains

$$\min\{\sum_{U\subseteq S} y_i(U)f_i(U) \mid y_i \in \mathbb{R}_+^{\mathcal{P}(S)}, \sum_{U\subseteq S} y_i(U)\chi^U = w_i\}. \tag{46.4}$$

So by Theorem 44.3, we can assume that the collections

$$\mathcal{F}_i := \{U \subseteq S \mid y_i(U) > 0\} \tag{46.5}$$

are chains. Hence, by Theorem 41.11, (46.2) has an optimum solution such that the inequalities with positive coefficients have a totally unimodular constraint matrix. Therefore, by Theorem 5.35, (46.1) is box-TDI. ∎

(This proof method is due to Edmonds [1970b].)

46.2. Consequences

Theorem 46.1 has the following consequences. First, the integrality of the intersection of two polymatroids:

Corollary 46.1a (polymatroid intersection theorem)**.** *The intersection of two integer (extended) polymatroids is box-integer.*

Proof. If P_{f_1} and P_{f_2} are integer polymatroids, f_1 and f_2 can be taken to be integer-valued, by Corollary 44.3g. Hence (46.1) determines a box-integer polyhedron. ∎

Next, a min-max relation:

Corollary 46.1b. *Let f_1 and f_2 be submodular set functions on S with $f_1(\emptyset) = f_2(\emptyset) = 0$. Then*

$$\max\{x(U) \mid x \in EP_{f_1} \cap EP_{f_2}\} = \min_{T\subseteq U}(f_1(T) + f_2(U \setminus T)) \tag{46.6}$$

for each $U \subseteq S$.

Proof. This follows by maximizing $w^{\mathsf{T}}x$ over (46.1) for $w := \chi^U$, and applying Theorem 46.1. ∎

Similarly, for (nonextended) polymatroids:

Corollary 46.1c. *Let f_1 and f_2 be nondecreasing submodular set functions on S with $f_1(\emptyset) = f_2(\emptyset) = 0$. Then*

$$\max\{x(U) \mid x \in P_{f_1} \cap P_{f_2}\} = \min_{T\subseteq U}(f_1(T) + f_2(U \setminus T)) \tag{46.7}$$

for each $U \subseteq S$.

Proof. As the previous corollary. ∎

Let f_1 and f_2 be submodular set functions on S with $f_1(\emptyset) = f_2(\emptyset) = 0$. Define

(46.8) $$f(U) := \min_{T \subseteq U}(f_1(T) + f_2(U \setminus T))$$

for $U \subseteq S$. It is easy to see that a vector x belongs to $P_{f_1} \cap P_{f_2}$ if and only if

(46.9) $$\begin{array}{ll} x_s \geq 0 & (s \in S), \\ x(U) \leq f(U) & (U \subseteq S). \end{array}$$

Moreover, system (46.9) is box-totally dual integral, since $f(U) \leq f_i(U)$ for each $U \subseteq S$ and $i = 1, 2$.

A consequence is that $P_{f_1} \cap P_{f_2}$ is integer if and only if f is integer. It may occur that P_{f_1} and P_{f_2} are not integer (i.e., f_1 and f_2 are not integer), while $P_{f_1} \cap P_{f_2}$ is integer (i.e., f is integer). For instance, take $P_{f_1} = \{(x_1, x_2) \mid 0 \leq x_1 \leq 1, 0 \leq x_2 \leq \frac{3}{2}\}$ and $P_{f_2} = \{(x_1, x_2) \mid (x_2, x_1) \in P_{f_1}\}$.

Many other results on polymatroid intersection may be deduced from Theorem 46.1, by considering derived polymatroids (cf. McDiarmid [1978]). For instance, if P_{f_1} and P_{f_2} are integer polymatroids in $\mathbb{R}^S$, v and w are integer vectors, and k and ℓ are integers, then the polytope

(46.10) $$\{x \in P_{f_1} \cap P_{f_2} \mid v \leq x \leq w, k \leq x(S) \leq \ell\}$$

is integer again. To see this, it suffices to show that the polytope $P_{f_1} \cap P_{f_2} \cap \{x \mid x(S) = k\}$ is integer for any integer k. We can reset $f_1(S) := \min\{f_1(S), k\}$. Then the polytope is a face of $P_{f_1} \cap P_{f_2}$, and hence is integer. In fact, the system determining (46.10) is box-TDI — see Corollary 49.12d.

The intersection of *three* integer polymatroids can have noninteger vertices, as the following example shows. Let $S = \{1, 2, 3, 4\}$ and let P_1, P_2 and P_3 be the following polymatroids:

(46.11) $$\begin{array}{l} P_1 := \{x \in \mathbb{R}^S \mid x \geq \mathbf{0}, x(\{1,2\}) \leq 1, x(\{3,4\}) \leq 1\}, \\ P_2 := \{x \in \mathbb{R}^S \mid x \geq \mathbf{0}, x(\{1,3\}) \leq 1, x(\{2,4\}) \leq 1\}, \\ P_3 := \{x \in \mathbb{R}^S \mid x \geq \mathbf{0}, x(\{1,4\}) \leq 1, x(\{2,3\}) \leq 1\}. \end{array}$$

(So each P_i is the independent set polytope of a partition matroid.) Now the vector $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2})$ is in $P_1 \cap P_2 \cap P_3$, but the only integer vectors in $P_1 \cap P_2 \cap P_3$ are the 0,1 vectors with at most one 1.

46.3. Contrapolymatroid intersection

Similar results as in the previous sections can be shown for the intersection of two contrapolymatroids. Such results can be proved similarly, or can be derived from the corresponding results for polymatroids.

Consider the system, for set functions g_1, g_2 on S:

(46.12) $$\begin{array}{ll} x(U) \geq g_1(U) & \text{for } U \subseteq S, \\ x(U) \geq g_2(U) & \text{for } U \subseteq S. \end{array}$$

Then Theorem 46.1 gives:

Corollary 46.1d. *If g_1 and g_2 are supermodular, then* (46.12) *is box-TDI.*

Proof. This follows from the box-total dual integrality of (46.1) taking $f_i := -g_i$ for $i = 1, 2$. ∎

46.4. Intersecting a polymatroid and a contrapolymatroid

Let S be a finite set. For set functions f and g on S consider the system

(46.13) $$\begin{aligned} x(U) &\leq f(U) \quad \text{for } U \subseteq S,\\ x(U) &\geq g(U) \quad \text{for } U \subseteq S. \end{aligned}$$

Theorem 46.2. *If f is submodular and g is supermodular, then system* (46.13) *is box-TDI.*

Proof. We can assume that $f(\emptyset) \geq 0$ and $g(\emptyset) \leq 0$. Choose $w \in \mathbb{R}^S$, and consider the dual problem of maximizing $w^{\mathsf{T}} x$ over (46.13):

(46.14) $$\begin{aligned} \min\{&\sum_{U \subseteq S} y(U) f(U) - \sum_{U \subseteq S} z(U) g(U) \mid \\ &y, z \in \mathbb{R}_+^{\mathcal{P}(S)}, \sum_{U \subseteq S} y(U)\chi^U - \sum_{U \subseteq S} z(U)\chi^U = w\}. \end{aligned}$$

Let y, z attain this minimum. Define

(46.15) $$u := \sum_{U \subseteq S} y(U)\chi^U \text{ and } v := \sum_{U \subseteq S} z(U)\chi^U.$$

So y attains

(46.16) $$\min\{\sum_{U \subseteq S} y(U) f(U) \mid y \in \mathbb{R}_+^{\mathcal{P}(S)}, \sum_{U \subseteq S} y(U)\chi^U = u\}$$

and z attains

(46.17) $$\max\{\sum_{U \subseteq S} z(U) g(U) \mid z \in \mathbb{R}_+^{\mathcal{P}(S)}, \sum_{U \subseteq S} z(U)\chi^U = v\}.$$

By Theorem 44.3, (46.16) has an optimum solution y with $\mathcal{F} := \{U \mid y(U) > 0\}$ is a chain. Similarly, (46.17) has an optimum solution z with $\mathcal{G} := \{U \mid z(U) > 0\}$ is a chain. Hence by Theorem 41.11, minimum (46.14) has an optimum solution such that the inequalities corresponding to positive coefficients have a totally unimodular constraint matrix. Hence by Theorem 5.35, (46.13) is box-TDI. ∎

So for the intersection of a polymatroid and a contrapolymatroid one gets:

Corollary 46.2a. *The intersection of an integer extended polymatroid and an integer extended contrapolymatroid is integer.*

Proof. Directly from the fact that an integer extended polymatroid is the solution set of $x(U) \leq f(U)$ $(U \subseteq S)$ for some integer submodular set function on S, and an integer extended contrapolymatroid is the solution set of $x(U) \geq g(U)$ $(U \subseteq S)$ for some integer submodular set function on S. Hence, by Theorem 46.2, the intersection is determined by a TDI system with integer right-hand sides. So the intersection is integer. ∎

46.5. Frank's discrete sandwich theorem

Frank [1982b] showed the following 'discrete sandwich theorem' (analogous to the 'continuous sandwich theorem', stating the existence of a linear function between a convex and a concave function):

Corollary 46.2b (Frank's discrete sandwich theorem)**.** *Let f be a submodular and g a supermodular set function on S, with $g \leq f$. Then there exists a modular set function h on S with $g \leq h \leq f$. If f and g are integer, h can be taken integer.*

Proof. We can assume that $g(\emptyset) = 0 = f(\emptyset)$, by resetting $f(U) := f(U) - f(\emptyset)$ and $g(U) := g(U) - f(\emptyset)$, for each $U \subseteq S$, and $g(\emptyset) := 0$.

Define $f'(U) := f(S) - g(S \setminus U)$ for each $U \subseteq S$. Then f' is submodular. Now by Corollary 46.1b:

(46.18) $$\max\{x(S) \mid x(U) \leq f(U), x(U) \leq f'(U) \text{ for each } U \subseteq S\} = \min\{f(T) + f'(S \setminus T) \mid T \subseteq S\}.$$

The minimum is at least $f(S)$, since $f(T) + f'(S \setminus T) = f(T) + f(S) - g(T) \geq f(S)$. Hence there exists an $x \in \mathbb{R}^S$ with $x(U) \leq f(U)$ and $x(U) \leq f'(U)$ for each $U \subseteq S$ and with $x(S) = f(S)$. Defining $h(U) := x(U)$, gives the modular function as required, since for each $U \subseteq S$:

(46.19) $$g(U) = f(S) - f'(S \setminus U) \leq x(S) - x(S \setminus U) = x(U) \leq f(U).$$

If f and g are integer, we can choose x integer, implying that h is integer. ∎

As Lovász [1983c] observed, the first part of this result can be derived from the continuous sandwich theorem: define $\tilde{f} : \mathbb{R}^S_+ \to \mathbb{R}$ by

(46.20) $$\tilde{f}(x) := \sum_{i=1}^{k} \lambda_i f(U_i),$$

where $\emptyset \neq U_1 \subset U_2 \subset \cdots \subset U_n \subseteq S$ and $\lambda_1, \ldots, \lambda_k > 0$ are such that $x = \sum_{i=1}^{k} \lambda_i \chi^{U_i}$. Define $\tilde{g}$ similarly. Then $\tilde{f}$ is convex and $\tilde{g}$ is concave, and $\tilde{g} \leq \tilde{f}$. Hence there is a linear function $\tilde{h}$ satisfying $\tilde{g} \leq \tilde{h} \leq \tilde{f}$. This gives the function h as required.

46.6. Integer decomposition

Integer polymatroids have the integer decomposition property. More generally:

Corollary 46.2c. *Let $P_1, \ldots, P_k$ be integer polymatroids. Then each integer vector in $P_1+\cdots+P_k$ is a sum $x_1+\cdots+x_k$ of integer vectors $x_1 \in P_1, \ldots, x_k \in P_k$.*

Proof. It suffices to show this for $k = 2$; the general case follows by induction (as the sum of integer polymatroids is again an integer polymatroid, by Theorem 44.6). Choose an integer vector $x \in P_1 + P_2$. Let Q be the contrapolymatroid given by

(46.21) $$Q := x - P_2.$$

Then $P_1 \cap Q \neq \emptyset$, since $x = x_1 + x_2$ for some $x_1 \in P_1$ and $x_2 \in P_2$, implying $x_1 \in P_1 \cap Q$. Now Q is integer as x and P_2 are integer. Hence by Corollary 46.2a, $P_1 \cap Q$ contains an integer vector y. Then $x - y \in P_2$, and so x is the sum of $y \in P_1$ and $x - y \in P_2$. ∎

This implies the integer decomposition property for integer polymatroids, proved by Giles [1975] (also by Baum and Trotter [1981]):

Corollary 46.2d. *An integer polymatroid has the integer decomposition property.*

Proof. Directly from Corollary 46.2c, by taking all P_i equal. ∎

This gives the following integer rounding properties (Baum and Trotter [1981]). Let P_f be the integer polymatroid determined by some integer submodular set function f on S. Let $\mathcal{B}$ be the collection of integer base vectors of P_f. Let B be the $\mathcal{B} \times S$ incidence matrix. Then for each $c \in \mathbb{Z}^S_+$, one has

(46.22) $$\min\{y^{\mathsf{T}}\mathbf{1} \mid y \in \mathbb{Z}^{\mathcal{B}}_+, y^{\mathsf{T}}B \geq c^{\mathsf{T}}\} = \lceil \min\{y^{\mathsf{T}}\mathbf{1} \mid y \in \mathbb{R}^{\mathcal{B}}_+, y^{\mathsf{T}}B \geq c^{\mathsf{T}}\} \rceil.$$

Indeed, $\geq$ is trivial. To see equality, let k be equal to the right-hand side. Then $c \in k \cdot P_f$, and hence, by Corollary 46.2d, $c \leq b_1 + \cdots + b_k$ for rows $b_1, \ldots, b_k$ of B. This shows equality.

Note that the right-hand side in (46.22) is equal to $\lceil \max\{c^{\mathsf{T}}x \mid x \in A(P_f)\} \rceil$, where $A(P_f)$ is the antiblocking polyhedron of P_f.

Similarly, one has:

(46.23) $$\max\{y^{\mathsf{T}}\mathbf{1} \mid y \in \mathbb{Z}^{\mathcal{B}}_+, y^{\mathsf{T}}B \leq c^{\mathsf{T}}\} = \lfloor \max\{y^{\mathsf{T}}\mathbf{1} \mid y \in \mathbb{R}^{\mathcal{B}}_+, y^{\mathsf{T}}B \leq c^{\mathsf{T}}\} \rfloor.$$

Now the right-hand side is equal to $\lfloor \min\{c^{\mathsf{T}}x \mid x \in B(Q)\}\rfloor$, where $B(Q)$ is the blocking polyhedron of $Q := \{x \in \mathbb{R}^S \mid x(U) \geq f(S) - f(S \setminus U)$ for $U \subseteq S\}$.

If f is the rank function of a matroid, then (46.22) describes the minimum number of bases covering S, while (46.23) describes the maximum number of disjoint bases.

46.7. Further results and notes

46.7a. Up and down hull of the common base vectors

Let f_1 and f_2 be nondecreasing submodular set functions on S, with $f_1(\emptyset) = f_2(\emptyset) = 0$ and $f_1(S) = f_2(S)$, and let P_1 and P_2 be the associated polymatroids. Let F_1 and F_2 be the faces of base vectors of P_1 and of P_2, respectively. Suppose that $F_1 \cap F_2 \neq \emptyset$, equivalently that

(46.24) $$f_1(S) = f_2(S) = \max\{x(S) \mid x \in P_1 \cap P_2\} = \min_{U \subseteq S} f_1(U) + f_2(S \setminus U).$$

Consider the polyhedra P and Q defined by

(46.25) $$P := \{x \in \mathbb{R}^S_+ \mid x \leq y \text{ for some } y \text{ in } F_1 \cap F_2\},$$
$$Q := \{x \in \mathbb{R}^S_+ \mid x \geq y \text{ for some } y \text{ in } F_1 \cap F_2\}.$$

So if f_1 and f_2 are the rank functions of matroids on S, then P is just the convex hull of incidence vectors of subsets of S which are contained in a common base.

Note that F_1 and F_2 are the faces of minimal vectors in the contrapolymatroids Q_1 and Q_2 associated with the supermodular set functions g_1 and g_2 on S given by

(46.26) $$g_i(U) := f_i(S) - f_i(S \setminus U)$$

for $U \subseteq S$ and $i = 1, 2$ (cf. Section 44.5). So $P \subseteq P_1 \cap P_2$ and $Q \subseteq Q_1 \cap Q_2$.

Let the set functions f and g on S be defined by

(46.27) $$f(U) := \max\{x(U) \mid x \in P_1 \cap P_2\} = \min_{T \subseteq U}(f_1(T) + f_2(U \setminus T)),$$
$$g(U) := \min\{x(U) \mid x \in Q_1 \cap Q_2\} = \max_{T \subseteq U}(g_1(T) + g_2(U \setminus T)),$$

for $U \subseteq S$ (cf. Corollary 46.1c). Then $f(S) = g(S) = f_1(S) = f_2(S) = g_1(S) = g_2(S)$.

It is easy to see that if x belongs to Q, then

(46.28) $$x(U) \geq f(S) - f(S \setminus U) \text{ for each } U \subseteq S$$

(note that $x \geq \mathbf{0}$ follows from (46.28) by taking $U = \{s\}$). Indeed, if $x \geq z$ with $z \in F_1 \cap F_2$, then $x(U) \geq z(U) = f(S) - z(S \setminus U) \geq f(S) - f(S \setminus U)$, as $z \in P_1 \cap P_2$.

Similarly, if x belongs to P, then

(46.29) $$x_s \geq 0 \quad (s \in S),$$
$$x(U) \leq g(S) - g(S \setminus U) \quad (U \subseteq S).$$

In fact, the systems (46.28) and (46.29) determine Q and P respectively. This was shown by Cunningham [1977] and McDiarmid [1978], thus proving a conjecture of Fulkerson [1971a] (cf. Weinberger [1976]).

Theorem 46.3. *Polyhedron Q is determined by* (46.28). *Polyhedron P is determined by* (46.29).

Proof. Consider $x \in \mathbb{R}^S_+$ and let P'_i be the polymatroid $P'_i := \{y \in P_i \mid y \leq x\}$ for $i = 1, 2$ (cf. Section 44.1). By (44.8), the submodular function f'_i associated with P'_i is given by

(46.30) $$f'_i(U) = \min_{T \subseteq U}(f_i(T) + x(U \setminus T))$$

for $U \subseteq S$ and $i = 1, 2$. Now x is in Q if and only if there is a vector z in $P_1 \cap P_2$ with $z \leq x$ and $z(S) = f(S)$, i.e., if and only if there is a vector z in $P'_1 \cap P'_2$ with $z(S) = f(S)$. By (46.7) such a vector exists if and only if

(46.31) $$\min_{U \subseteq S}(f'_1(U) + f'_2(S \setminus U)) \geq f(S).$$

Substituting (46.30) one finds that (46.31) is equivalent to (46.28).

The second statement of Theorem 46.3 is proved similarly. ∎

This theorem has a self-refining character. If k is a rational number with $k \leq f(S)$ and if $w \in Q$, then

(46.32) $$\begin{aligned}&\{x \in \mathbb{R}^S_+ \mid x \geq z \text{ for some } z \text{ in } P_1 \cap P_2 \text{ with } z(S) = k\}\\ &= \{x \in \mathbb{R}^S_+ \mid x(U) \geq k - f(S \setminus U) \text{ for all } U \subseteq S\}\end{aligned}$$

and

(46.33) $$\begin{aligned}&\{x \in \mathbb{R}^S_+ \mid x \geq z \text{ for some } z \text{ in } F_1 \cap F_2 \text{ with } z \leq w\}\\ &= \{x \in \mathbb{R}^S_+ \mid x(S \setminus (T \cup U)) \geq f(S) - w(U) - f(T) \text{ for disjoint}\\ &\qquad T, U \subseteq S\},\end{aligned}$$

as can be seen by taking appropriate subpolymatroids of P_1 and P_2 (cf. also McDiarmid [1976,1978]).

This has the following applications. Let $G = (V, E)$ be a bipartite graph, let $x \in \mathbb{R}^E_+$, and let k be a natural number. Then there exists a vector $z \leq x$ such that z is a convex combination of incidence vectors of matchings in G of size k if and only if

(46.34) $$x(E[U]) \geq k - |V| + |U|$$

for all $U \subseteq V$ (where $E[U]$ denotes the set of edges spanned by U). This can be derived as follows. Let V_1 and V_2 be the colour classes of G. For $F \subseteq E$, let $f_i(F)$ be the number of vertices in V_i covered by F (for $i = 1, 2$). Then $f(F)$ equals the maximum size of a matching in F, which is equal to the minimum number of vertices covering F. Hence the inequalities $x(F) \geq k - f(E \setminus F)$ (for $F \subseteq E$) follow from $x(E[U]) \geq k - |V \setminus U|$ (for $U \subseteq V$).

Another application is Corollary 52.3a on the up hull of the r-arborescence polytope (cf. Section 52.1a).

Gröflin and Hoffman [1981] gave a method to show the following:

Theorem 46.4. (46.28) *and* (46.29) *are box-TDI.*

Proof. We prove that (46.28) is box-TDI. The box-total dual integrality of (46.29) is proved similarly.

Let $\mathcal{R}$ be the collection of all pairs (T, U) of subsets of S with $T \cap U = \emptyset$. Then the system

$$(46.35) \qquad x(S \setminus (T \cup U)) \geq f(S) - f_1(T) - f_2(U) \qquad ((T,U) \in \mathcal{R})$$

is equivalent to (46.28), in the following sense: by (46.27), (46.35) determines Q, and (46.35) contains all inequalities occurring in (46.28); moreover, all inequalities in (46.35) satisfied with equality by some $x \in Q$, also occur in (46.28). Hence, if (46.35) is box-totally dual integral, also (46.28) is box-totally dual integral. So it suffices to show the box-total dual integrality of (46.35). To this end, let $w \in \mathbb{Z}_+^S$, and consider the dual of minimizing $w^\mathsf{T} x$ over (46.35):

$$(46.36) \qquad \max\{\sum_{(T,U)\in\mathcal{R}} y(T,U)(f(S) - f_1(T) - f_2(U)) \mid y \in \mathbb{R}_+^{\mathcal{R}}, \sum_{(T,U)\in\mathcal{R}} y(T,U)\chi^{S\setminus(T\cup U)} = w\}.$$

We show that it is attained by an integer vector y if w is integer.

To this end, let y attain the maximum (46.36) such that

$$(46.37) \qquad \sum_{(T,U)\in\mathcal{R}} y(T,U)(|T| + |S \setminus U|)(|U| + |S \setminus T|)$$

is as small as possible. Then:

(46.38) if $y(A, B) > 0$ and $y(C, D) > 0$, then either $A \subseteq C$ and $B \supseteq D$, or $A \supseteq C$ and $B \subseteq D$.

Suppose not. Define $\alpha := \min\{y(A, B), y(C, D)\}$. Define $y' : \mathcal{R} \to \mathbb{R}_+$ by

$$(46.39) \qquad \begin{aligned} y'(A,B) &:= y(A,B) - \alpha, \\ y'(C,D) &:= y(C,D) - \alpha, \\ y'(A\cap C, B\cup D) &:= y(A\cap C, B\cup D) + \alpha, \\ y'(A\cup C, B\cap D) &:= y(A\cup C, B\cap D) + \alpha, \end{aligned}$$

and let y' coincide with y in the other components. One easily checks that

$$(46.40) \qquad \begin{aligned} &\sum_{(T,U)\in\mathcal{R}} y'(T,U)\chi^{S\setminus(T\cup U)} = \sum_{(T,U)\in\mathcal{R}} y(T,U)\chi^{S\setminus(T\cup U)} \text{ and} \\ &\sum_{(T,U)\in\mathcal{R}} y'(T,U)(f(S) - f_1(T) - f_2(U)) \\ &\geq \sum_{(T,U)\in\mathcal{R}} y(T,U)(f(S) - f_1(T) - f_2(U)), \end{aligned}$$

by the submodularity of f_1 and f_2. So y' also attains the maximum (46.36). Moreover, one straightforwardly checks that replacing y by y' decreases (46.37).[36] This contradicts our assumption, and therefore proves (46.38).

[36] This can be seen with Theorem 2.1: Make a copy $\widetilde{S}$ of S, and, for any $U \subseteq S$, let $\widetilde{U}$ be the set of copies of elements of U. Define $X_{T,U} := T\cup(\widetilde{S}\setminus\widetilde{U})$. Then $|T|+|S\setminus U| = |X_{T,U}|$ and $|U| + |S \setminus T| = |(S \cup \widetilde{S}) \setminus X_{T,U}|$. Moreover, for (A, B) and (C, D) in $\mathcal{R}$ we have $X_{A\cap C,B\cup D} = X_{A,B} \cap X_{C,D}$ and $X_{A\cup C,B\cap D} = X_{A,B} \cup X_{C,D}$. So the replacements decrease (46.37) by Theorem 2.1, since $X_{A,B} \not\subseteq X_{C,D} \not\subseteq X_{A,B}$.

Let $\mathcal{R}_0 := \{(T,U) \in \mathcal{R} \mid y(T,U) > 0\} = \{(T_1,U_1),\ldots,(T_n,U_n)\}$ with $T_1 \subseteq \cdots \subseteq T_n$ and $U_n \subseteq \cdots \subseteq U_1$ (this is possible by (46.38)). Let M be the $\{0,1\}$ matrix with rows indexed by $\mathcal{R}_0$ and columns indexed by S, such that $M_{(T,U),s} = 1$ if and only if $s \notin T \cup U$. Then for each s in S, the indices i for which $M_{(T_i,U_i),s} = 1$ form a contiguous interval of $\{1,\ldots,n\}$. Hence M is totally unimodular (as it is a network matrix with directed tree being a directed path). So we have the box-total dual integrality of (46.35) by Theorem 5.35. ∎

Frank and Tardos [1984a] indicated a direct derivation of this theorem from the total dual integrality of (46.1).

There are a number of straightforward corollaries. As for the integrality of polyhedra:

Corollary 46.4a. *If f (or, equivalently, g) is integer, then the polyhedra P, Q, and $F_1 \cap F_2$ are integer.*

Proof. This follows directly from Theorem 46.4. Note that $F_1 \cap F_2$ is integer if and only if P is integer. ∎

Also a min-max relation follows:

Corollary 46.4b. *Let $M_1 = (S,\mathcal{I}_1)$ and $M_2 = (S,\mathcal{I}_2)$ be matroids, with rank functions r_1 and r_2 and let k be the maximum size of a common independent set. Then for any subset U of S,*

(46.41) $$\min_I |U \cap I| = \max_{S_1,\ldots,S_t} \sum_{i=1}^{t} (k - r(S \setminus S_i)),$$

where the minimum ranges over all common independent sets I with $|I| = k$, and where the maximum ranges over all partitions of U into sets $S_1,\ldots,S_t$ ($t \geq 0$), and where $r(T)$ denotes the maximum size of a common independent set contained in T.

Proof. Apply Theorem 46.4, taking $c := \chi^U$, $f_i := r_i$, and $f := r$. ∎

It is not necessarily true that if $F_1 \cap F_2$ is integer, then also $P_1 \cap P_2$ (or $Q_1 \cap Q_2$) is integer — i.e., that the converse implication of Corollary 46.4a holds. For instance, if

(46.42) $$P_1 := \{(x,y,z)^{\mathsf{T}} \mid 0 \leq x \leq 1, 0 \leq y \leq 1, 0 \leq z \leq \tfrac{3}{2}, x+z \leq 2\},$$
$$P_2 := \{(x,y,z)^{\mathsf{T}} \mid 0 \leq x \leq 1, 0 \leq y \leq 1, 0 \leq z \leq \tfrac{3}{2}, y+z \leq 2\},$$

then $F_1 \cap F_2 = \{(1,1,1)^{\mathsf{T}}\}$, but $(\frac{1}{2},\frac{1}{2},\frac{3}{2})^{\mathsf{T}}$ is a vertex of $P_1 \cap P_2$.

Related results on integer decomposition of integer polymatroids in McDiarmid [1983].

46.7b. Further notes

Giles [1975] characterized the facets of the intersection of two polymatroids. Ageev [1988] studied the problem of maximizing a concave function over the intersection of polymatroids.

Chapter 47

Polymatroid intersection algorithmically

In this chapter we consider the problem of finding a vector of maximum weight in the intersection of two (extended) polymatroids algorithmically. We describe a strongly polynomial-time algorithm for this problem in four stages (where f_1 and f_2 are submodular set functions on S):

- a strongly polynomial-time algorithm finding a maximum-size vector in $EP_{f_1} \cap EP_{f_2}$ (Section 47.1),
- a strongly polynomial-time algorithm finding a common base vector of f_1 and f_2 maximizing $x(s)$ for some prescribed $s \in S$ (Section 47.2),
- a polynomial-time algorithm finding a maximum-weight common base vector of f_1 and f_2 (Section 47.3),
- a strongly polynomial-time algorithm finding a maximum-weight common base vector of f_1 and f_2 (Section 47.4).

At the base of the algorithms is submodular function minimization, which leads back to the 'consistent breadth-first search' technique proposed in a pioneering paper of Schönsleben [1980] on polymatroid intersection.

47.1. A maximum-size common vector in two polymatroids

We first consider the problem:

(47.1) given: submodular set functions f_1 and f_2 on a set S (by value giving oracles),
find: an $x \in EP_{f_1} \cap EP_{f_2}$ maximizing $x(S)$, and a $T \subseteq S$ with $x(S) = f_1(T) + f_2(S \setminus T)$.

So T certifies that x indeed maximizes $x(S)$ over $EP_{f_1} \cap EP_{f_2}$, since for any $x' \in EP_{f_1} \cap EP_{f_2}$ we have:

(47.2) $$x'(S) = x'(T) + x'(S \setminus T) \leq f_1(T) + f_2(S \setminus T) = x(S).$$

On the other hand, x certifies that T minimizes $f_1(T) + f_2(S \setminus T)$.

Then (Lawler and Martel [1982a], extending a weakly polynomial bound of Schönsleben [1980]):

Theorem 47.1. *Problem* (47.1) *is solvable in strongly polynomial-time.*

Proof. We can assume that $f_1(\emptyset) = 0$ and $f_2(\emptyset) = 0$. Define the submodular set function f on S by

$$f(U) := f_1(U) + f_2(S \setminus U) - f_2(S) \tag{47.3}$$

for $U \subseteq S$. With the submodular function minimization algorithm described in Section 45.4 we find a subset T of S minimizing f. So, by Corollary 46.1b, $f(T) + f_2(S)$ is equal to the maximum of $x(S)$ over $EP_{f_1} \cap EP_{f_2}$.

The submodular function minimization algorithm of Section 45.4 also gives vertices $b_1, \ldots, b_k$ of EP_f and $\lambda_1, \ldots, \lambda_k \geq 0$ with $\lambda_1 + \cdots + \lambda_k = 1$ such that for

$$y := \lambda_1 b_1 + \cdots + \lambda_k b_k \tag{47.4}$$

we have $y(T) = f(T)$, $\text{supp}^-(y) \subseteq T$, and $\text{supp}^+(y) \subseteq S \setminus T$. (Here, as usual, $\text{supp}^+(x) := \{s \in S \mid x(s) > 0\}$ and $\text{supp}^-(x) := \{s \in S \mid x(s) < 0\}$.)

Now for each $i = 1, \ldots, k$, we can find $b'_i \in EP_{f_1}$ and $b''_i \in EP_{f_2}$ with $b_i = b'_i - b''_i$. Indeed, let $u_1, \ldots, u_n$ be a total order of S generating b_i. (That is, $b_i(u_j) = f(\{u_1, \ldots, u_j\}) - f(\{u_1, \ldots, u_{j-1}\})$ for $j = 1, \ldots, n$. These orderings are also implied by the submodular function minimization algorithm.) Let b'_i be the vertex of EP_{f_1} generated by the order $u_1, \ldots, u_n$ (that is, $b'_i(u_j) = f_1(\{u_1, \ldots, u_j\}) - f_1(\{u_1, \ldots, u_{j-1}\})$ for $j = 1, \ldots, n$). Let b''_i be the vertex of EP_{f_2} generated by the order $u_n, u_{n-1}, \ldots, u_1$ (that is, $b''_i(u_j) = f_2(\{u_j, \ldots, u_n\}) - f_2(\{u_{j+1}, \ldots, u_n\})$ for $j = 1, \ldots, n$). Then by definition of f we have $b_i = b'_i - b''_i$, since for each j:

$$\begin{aligned} b_i(u_j) &= f(\{u_1, \ldots, u_j\}) - f(\{u_1, \ldots, u_{j-1}\}) \\ &= f_1(\{u_1, \ldots, u_j\}) + f_2(\{u_{j+1}, \ldots, u_n\}) - f_1(\{u_1, \ldots, u_{j-1}\}) \\ &\quad - f_2(\{u_j, \ldots, u_n\}) = b'_i(u_j) - b''_i(u_j). \end{aligned} \tag{47.5}$$

Define

$$x' := \sum_{i=1}^{k} \lambda_i b'_i, \; x'' := \sum_{i=1}^{k} \lambda_i b''_i, \text{ and } x := x' \wedge x'', \tag{47.6}$$

where $\wedge$ stands for taking coordinatewise minimum. As $x' \in EP_{f_1}$ and $x'' \in EP_{f_2}$, we know $x \in EP_{f_1} \cap EP_{f_2}$. Also, as $y = x' - x''$, we know that if $u \in T$, then $y(u) \leq 0$, hence $x''(u) \geq x'(u)$, and therefore $x(u) = x'(u)$. Similarly, if $u \in S \setminus T$, then $x(u) = x''(u)$. Hence

$$\begin{aligned} x(S) &= x(T) + x(S \setminus T) = x'(T) + x''(S \setminus T) = (x' - x'')(T) + x''(S) \\ &= y(T) + x''(S) = f(T) + f_2(S) = f_1(T) + f_2(S \setminus T), \end{aligned} \tag{47.7}$$

as required. ∎

47.2. Maximizing a coordinate of a common base vector

Theorem 47.1 implies the strong polynomial-time solvability of:

(47.8) given: submodular set functions f_1 and f_2 on a set S (by value giving oracles) and an element $s \in S$,
find: a common base vector x of f_1 and f_2 maximizing $x(s)$, and subsets S_1 and S_2 of S with $S_1 \cap S_2 = \{s\}$, $S_1 \cup S_2 = S$, and $x(S_i) = f_i(S_i)$ for $i = 1, 2$.

This is a result of Frank [1984a]:

Theorem 47.2. *Problem* (47.8) *is solvable in strongly polynomial time.*

Proof. We can assume that $f_1(S) = f_2(S)$ and that f_1 and f_2 have a common base vector (this can be checked by Theorem 47.1). Hence

(47.9) $$f_1(S) \leq f_1(U) + f_2(S \setminus U)$$

for each $U \subseteq S$. Define $S' := S \setminus \{s\}$.

First determine S_1, S_2 with $S_1 \cap S_2 = \{s\}$ and $S_1 \cup S_2 = S$ and minimizing $f_1(S_1) + f_2(S_2)$. This can be done by minimizing the submodular function $f_1(U + s) + f_2(S \setminus U)$ over $U \subseteq S'$.

Define

(47.10) $$\alpha := f_1(S_1) + f_2(S_2) - f_1(S).$$

For $i = 1, 2$ and $U \subseteq S'$, define

(47.11) $$g_i(U) := \min\{f_i(U), f_i(U + s) - \alpha\}.$$

Then g_1 and g_2 are submodular set functions on S', as is easy to check. Moreover,

(47.12) $$g_i(S') = f_i(S) - \alpha.$$

To show this, we may assume that $i = 1$. Then we must show:

(47.13) $$f_1(S') \geq f_1(S) - \alpha = 2f_1(S) - f_1(S_1) - f_2(S_2).$$

Now $f_1(S_1 \setminus \{s\}) + f_2(S_2) \geq f_1(S)$ (since f_1 and f_2 have a common base vector) and $f_1(S_1) - f_1(S_1 \setminus \{s\}) \geq f_1(S) - f_1(S')$ (by the submodularity of f_1). These two inequalities imply (47.13).

Then

(47.14) g_1 and g_2 have a common base vector.

Otherwise, S' can be partitioned into sets R_1 and R_2 with

(47.15) $$g_1(R_1) + g_2(R_2) < g_1(S').$$

If $g_1(R_1) = f_1(R_1)$ and $g_2(R_2) = f_2(R_2)$, then (47.15) implies

(47.16) $$2f_1(S) > f_1(S_1) + f_2(S_2) + f_1(R_1) + f_2(R_2) \geq f_1(S_1 \cap R_1) + f_1(S_1 \cup R_1) + f_2(S_2 \cap R_2) + f_2(S_2 \cup R_2).$$

By symmetry, we can assume that $f_1(S) > f_1(S_1 \cap R_1) + f_2(S_2 \cup R_2)$. However, $S_1 \cap R_1$ and $S_2 \cup R_2$ partition S, contradicting (47.9).

If $g_1(R_1) = f_1(R_1)$ and $g_2(R_2) = f_2(R_2 + s) - \alpha$, then (47.15) implies

(47.17) $$f_1(S) - \alpha > f_1(R_1) + f_2(R_2 + s) - \alpha,$$

and hence $f_1(S) > f_1(R_1) + f_2(R_2 + s)$, contradicting (47.9).

If $g_1(R_1) = f_1(R_1 + s) - \alpha$ and $g_2(R_2) = f_2(R_2 + s) - \alpha$, then (47.15) implies

(47.18) $$f_1(S) - \alpha > f_1(R_1 + s) + f_2(R_2 + s) - 2\alpha,$$

implying $f_1(S_1) + f_2(S_2) > f_1(R_1 + s) + f_2(R_2 + s)$, contradicting the minimality of $f_1(S_1) + f_2(S_2)$. This proves (47.14).

By Theorem 47.1, we can find in strongly polynomial time a common base vector x of g_1 and g_2. So $x(S') = g_1(S')$. Extend x to S by defining $x(s) := \alpha$. Then

(47.19) x is a common base vector of f_1 and f_2.

By symmetry, it suffices to show that x is a base vector of f_1. First, x belongs to EP_{f_1}, since for each $U \subseteq S'$ we have

(47.20) $$x(U) \leq g_1(U) \leq f_1(U) \text{ and } x(U + s) = x(U) + \alpha \leq g_1(U) + \alpha \leq f_1(U + s).$$

Next, x is a base vector of f_1, since

(47.21) $$x(S) = x(S') + \alpha = g_1(S') + \alpha = f_1(S),$$

by (47.12). This proves (47.19).

Moreover,

(47.22) $$x(S_i) = f_i(S_i)$$

for $i = 1, 2$. Indeed (for $i = 1$),

(47.23) $$x(S_1) = x(S) - x(S_2 \setminus \{s\}) \geq f_1(S) - g_2(S_2 \setminus \{s\}) \geq f_1(S) - f_2(S_2) + \alpha = f_1(S_1).$$

This proves (47.22), which implies that x is a common base vector of f_1 and f_2 maximizing $x(s)$, as for any common base vector z of f_1 and f_2 we have

(47.24) $$z(s) = z(S_1) + z(S_2) - z(S) \leq f_1(S_1) + f_2(S_2) - f_1(S) = x(S_1) + x(S_2) - x(S) = x(s).$$

So x, S_1, and S_2 have the required properties. ∎

47.3. Weighted polymatroid intersection in polynomial time

It may be shown with the ellipsoid method that the following problem is solvable in polynomial time:

(47.25) given: submodular functions f_1, f_2 on a set S (by value giving oracles) and a function $w : S \to \mathbb{Z}$,
find: a common base vector x of f_1 and f_2 maximizing $w^{\mathsf{T}}x$, and $w_1, w_2 : S \to \mathbb{Z}$ with $w = w_1 + w_2$ such that, for each $i = 1, 2$, x maximizes $w_i^{\mathsf{T}}x$ over all base vectors of f_i.

Cunningham and Frank [1985] gave, with the help of Theorem 47.2, a combinatorial polynomial-time algorithm (using submodular function minimization).

In order to describe this, we first give an auxiliary result concerning polymatroids. Let f be a submodular set function on S and let F be a face of EP_f. Define

(47.26) $$F^{\downarrow} := F - \mathbb{R}_+^S.$$

Then $F^{\downarrow}$ is an extended polymatroid again. Moreover, algorithmic properties for $F^{\downarrow}$ can be deduced from those for EP_f:

Lemma 47.3α. *Let f be a submodular set function on S, let $w : S \to \mathbb{Z}_+$, and let F be the set of vectors x in EP_f maximizing $w^{\mathsf{T}}x$. Then there is a submodular set function f' on S with $F^{\downarrow} = EP_{f'}$. Moreover, if f is given by a value giving oracle, $f'(U)$ can be computed in strongly polynomial time, for any $U \subseteq S$.*

Proof. We can assume that $f(\emptyset) = 0$. Let $\emptyset \neq T_1 \subset T_2 \subset \cdots \subset T_{k-1} \subset T_k = S$ be the (unique) sets satisfying

(47.27) $$w = \lambda_1 \chi^{T_1} + \cdots + \lambda_k \chi^{T_k},$$

for some $\lambda_1, \ldots, \lambda_{k-1} > 0$ and $\lambda_k \geq 0$. Set $T_0 := \emptyset$, and define f' by:

(47.28) $$f'(U) := \sum_{i=1}^{k} (f((U \cap T_i) \cup T_{i-1}) - f(T_{i-1})),$$

for $U \subseteq S$. Then f' is submodular, as it is the sum of k submodular functions. Also,

(47.29) $$f' \leq f,$$

since for each U we have, by the submodularity of f:

(47.30) $$\begin{aligned} f'(U) &= \sum_{i=1}^{k} (f((U \cap T_i) \cup T_{i-1}) - f(T_{i-1})) \\ &\leq \sum_{i=1}^{k} (f(U \cap T_i) - f(U \cap T_{i-1})) = f(U). \end{aligned}$$

We show:

(47.31) $\quad F^{\downarrow} = EP_{f'}.$

To see $\subseteq$, it suffices to show that $F \subseteq EP_{f'}$. Let $x \in F$. So $x(T_i) = f(T_i)$ for $i = 0, \ldots, k-1$. Hence

(47.32)
$$x(U) = \sum_{i=1}^{k} x(U \cap (T_i \setminus T_{i-1})) = \sum_{i=1}^{k} (x((U \cap T_i) \cup T_{i-1}) - x(T_{i-1}))$$
$$\leq \sum_{i=1}^{k} (f((U \cap T_i) \cup T_{i-1}) - f(T_{i-1})) = f'(U)$$

for each $U \subseteq S$. This proves that $x \in EP_{f'}$.

To see $\supseteq$ in (47.31), it suffices to show that any $x \in EP_{f'}$ with $x(S) = f'(S)$ belongs to F. As $f' \leq f$ we know that $x \in EP_f$. So it suffices to show that $x(T_j) = f(T_j)$ for $j = 1, \ldots, k$ (as this implies that x maximizes $w^{\mathsf{T}}x$ over EP_f, by the greedy algorithm). Now, as $f'(S) = f(S)$:

(47.33)
$$x(T_j) = x(S) - x(S \setminus T_j) \geq f'(S) - f'(S \setminus T_j)$$
$$= f(S) - \sum_{i=1}^{k} (f((T_i \setminus T_j) \cup T_{i-1}) - f(T_{i-1}))$$
$$= f(S) - \sum_{i=j+1}^{k} (f(T_i) - f(T_{i-1})) = f(T_j).$$

This proves (47.31). ∎

We also will use the following lemma:

Lemma 47.3β. *Let f be a submodular set function on S, let $w : S \to \mathbb{Z}$, and let F be the set of base vectors x of f maximizing $w^{\mathsf{T}}x$. Let $U \subseteq S$ and let x maximize $x(U)$ over F. Then x maximizes $(w + \chi^U)^{\mathsf{T}}x$ over all base vectors of f.*

Proof. Let $w' := w + \chi^U$. As x maximizes $x(U)$ over F, we know that x maximizes ${w'}^{\mathsf{T}}x$ over F. Also, some $z \in F$ maximizes ${w'}^{\mathsf{T}}z$ over EP_f, by the greedy method, as there is an ordering of S in which both w and w' are monotonically nondecreasing, and so EP_f has a vertex z maximizing both $w^{\mathsf{T}}z$ and ${w'}^{\mathsf{T}}z$.

As ${w'}^{\mathsf{T}}x \geq {w'}^{\mathsf{T}}z$, x maximizes ${w'}^{\mathsf{T}}x$ over EP_f. ∎

Now we can derive:

Theorem 47.3. *Problem* (47.25) *is solvable in polynomial time.*

Proof. We give a polynomial-time algorithm to transform a solution of (47.25) for some w to a solution of (47.25) for $w := w + \chi^s$, for any $s \in S$. This implies a polynomial-time algorithm for (47.25), since we can assume that $w \geq \mathbf{0}$, and since any $w \geq \mathbf{0}$ can be obtained from $w = \mathbf{0}$ by a polynomially bounded number of resettings $w := 2w$ and $w := w + \chi^s$ for $s \in S$. Note that for $w = \mathbf{0}$, (47.25) is trivial, and that a solution x, w_1, w_2 for w yields a solution $x, 2w_1, 2w_2$ for $2w$.

Let $s \in S$. Let x, w_1, w_2 be a solution of (47.25) for some w. For $i = 1, 2$, let F_i be the set of all vectors $x \in EP_{f_i}$ maximizing $w_i^{\mathsf{T}} x$ and let f_i' be a submodular function satisfying $F_i^{\downarrow} = EP_{f_i'}$ (Lemma 47.3α). Applying Theorem 47.2 to f_1', f_2', we find a common base vector x' of f_1' and f_2' maximizing $x'(s)$, and subsets S_1, S_2 of S with $S_1 \cap S_2 = \{s\}$, $S_1 \cup S_2 = S$, and $x'(S_1) = f_1'(S_1)$, $x'(S_2) = f_2'(S_2)$. Then x' maximizes $x'(S_1)$ over $EP_{f_1'}$, and x' maximizes $x'(S_2)$ over $EP_{f_2'}$. Hence, by Lemma 47.3β, x' is a base vector of f_1' maximizing $(w_1 + \chi^{S_1})^{\mathsf{T}} x'$, and also, x' is a base vector of f_2' maximizing $(w_2 + \chi^{S_2})^{\mathsf{T}} x'$. So

(47.34) $$x', w_1' := w_1 + \chi^{S_1} - \chi^S, w_2' := w_2 + \chi^{S_2},$$

gives a solution of (47.25) for $w + \chi^s$. ∎

47.4. Weighted polymatroid intersection in strongly polynomial time

A general simultaneous diophantine approximation method of Frank and Tardos [1985,1987] implies that (47.25) is *strongly* polynomial-time solvable. Fujishige, Röck, and Zimmermann [1989] showed that from Theorem 47.3 a combinatorial strongly polynomial-time algorithm can be derived, by extending the method of Tardos [1985a] for the minimum-cost circulation problem.

To prove this, we first show a sensitivity result. Let f_1, f_2 be submodular set functions on S. Call a pair $w_1, w_2 : S \to \mathbb{R}$ *good* if there exists an x that maximizes $w_i^{\mathsf{T}} x$ over EP_{f_i}, for both $i = 1$ and $i = 2$.

Lemma 47.4α. *Let $w : S \to \mathbb{Q}$ and let w_1, w_2 be a good pair with $w = w_1 + w_2$. Then for any $\tilde{w} : S \to \mathbb{Q}$ with $\tilde{w} \geq w$ there exists a good pair $\tilde{w}_1, \tilde{w}_2$ with $\tilde{w} = \tilde{w}_1 + \tilde{w}_2$ and $\|\tilde{w}_i - w_i\|_\infty \leq \|\tilde{w} - w\|_1$ for $i = 1, 2$.*

Proof. We can assume that w and $\tilde{w}$ are integer, and that $\|\tilde{w} - w\|_1 = 1$ (as the general case then follows inductively).

Let F_i be the set of all x maximizing $w_i^{\mathsf{T}} x$ over EP_{f_i}. Let f_i' be a submodular function satisfying $F_i^{\downarrow} = EP_{f_i'}$. Let s be such that $\tilde{w}(s) = w(s) + 1$. By the solvability of problem (47.8), there is a common base vector x of f_1' and f_2' maximizing x_s, and there exist S_1, S_2 with $S_1 \cap S_2 = \{s\}$ and $S_1 \cup S_2 = S$ such that $x(S_1) = f_1'(S_1)$ and $x(S_2) = f_2'(S_2)$. Define

(47.35) $$\tilde{w}_1 := w_1 + \chi^{S_1} - \chi^S, \tilde{w}_2 := w_2 + \chi^{S_2}.$$

So $\tilde{w} = \tilde{w}_1 + \tilde{w}_2$. By Lemma 47.3$\beta$, x maximizes $\tilde{w}_i^\mathsf{T} x$ over EP_{f_i} for $i = 1, 2$. Therefore, the pair $\tilde{w}_1, \tilde{w}_2$ is good. As $\|\tilde{w}_i - w_i\|_\infty \leq 1$ for $i = 1, 2$, this proves the lemma. ∎

Theorem 47.4. *Given submodular functions f_1, f_2 on a set S and $w \in \mathbb{Q}^S$, one can find a common base vector x of f_1 and f_2 maximizing $w^\mathsf{T} x$, in strongly polynomial time.*

Proof. Let be given submodular functions f_1, f_2 on a set S and a function $w : S \to \mathbb{Q}$. We may assume that f_1 and f_2 have a common base vector. (This can be checked by Theorem 47.1.)

We keep chains $\mathcal{C}_1, \mathcal{C}_2$ of subsets of S such that for $i = 1, 2$ and each $U \in \mathcal{C}_i$:

(47.36) $x(U) = f_i(U)$ for each common base vector x of f_1 and f_2 maximizing $w^\mathsf{T} x$,

and such that $S \in \mathcal{C}_1$ and $S \in \mathcal{C}_2$. Initially we set $\mathcal{C}_i := \{S\}$ for $i = 1, 2$. We describe an iteration that either extends $\mathcal{C}_1$ or $\mathcal{C}_2$, or gives a solution x.

We can assume that, for $i = 1, 2$,

(47.37) each base vector x of f_i satisfies $x(U) = f_i(U)$ for each $U \in \mathcal{C}_i$.

Indeed, let F_i be the set of vectors x in EP_{f_i} with $x(U) = f_i(U)$ for each $U \in \mathcal{C}_i$. So F_i is equal to the set of $x \in EP_{f_i}$ maximizing $c_i^\mathsf{T} x$, where $c_i := \sum_{U \in \mathcal{C}_i} \chi^U$. By Lemma 47.3$\alpha$, we can find f_i' with $F_i^\downarrow = EP_{f_i'}$. By (47.36), replacing the f_i by f_i' does not change the set of optimum solutions x of our problem.

Let

(47.38) $L :=$ linear hull of $\{\chi^U \mid U \in \mathcal{C}_1 \cup \mathcal{C}_2\}$.

Determine $y \in L$ minimizing

(47.39) $\|w - y\|_\infty$.

This can be done in strongly polynomial time as follows. For $i = 1, 2$, let $\mathcal{P}_i$ be the partition of S into nonempty classes such that u and v belong to the same class if and only if there is no set in $\mathcal{C}_i$ containing exactly one of u, v. Let D be the directed graph with vertex set $\mathcal{P}_1 \cup \mathcal{P}_2$ such that for each $v \in S$ there is an arc of length $w(v)$ from $U \in \mathcal{P}_1$ to $W \in \mathcal{P}_2$ with $v \in U \cap W$, and an arc of length $-w(v)$ in the reverse direction. Determine the minimum mean-length α of a directed circuit in D (cf. Section 8.5). It is the minimum α for which there exist $p_i : \mathcal{P}_i \to \mathbb{Q}$ such that

(47.40) $-\alpha \leq w(v) + p_1(U) - p_2(W) \leq \alpha$

for each arc as described. Then

(47.41) $$y := -\sum_{U \in \mathcal{P}_1} p_1(U) \chi^U + \sum_{W \in \mathcal{P}_2} p_2(W) \chi^W$$

minimizes (47.39).

Let α be the value of (47.39). If $\alpha = 0$, then $w \in L$, and so

(47.42) $$w = \sum_{i=1}^{2} \sum_{U \in \mathcal{C}_i} \lambda_i(U)\chi^U$$

for functions $\lambda_i : \mathcal{C}_i \to \mathbb{Q}$. Then for any common base vector x of f_1 and f_2 we have

(47.43) $$w^{\mathsf{T}}x = \sum_{i=1}^{2} \sum_{U \in \mathcal{C}_i} \lambda_i(U)x(U) = \sum_{i=1}^{2} \sum_{U \in \mathcal{C}_i} \lambda_i(U)f_i(U).$$

So each common base vector is optimum. As we can find any common base vector in strongly polynomial time (by Theorem 47.1), we have solved the problem.

So we can assume that $\alpha > 0$. Define $w' : S \to \mathbb{Z}$ by

(47.44) $$w' := \lfloor \frac{5n^2}{\alpha}(w - y) \rfloor,$$

where $n := |S|$. By definition of α, $\|w'\|_\infty = 5n^2$. Hence by Theorem 47.3, we can find in strongly polynomial time a common base vector x' of f_1 and f_2 and $w'_1, w'_2 : S \to \mathbb{Z}$ with $w' = w'_1 + w'_2$ such that x' is a base vector of f_i maximizing ${w'}_i^{\mathsf{T}}x'$, for $i = 1, 2$.

For $i = 1, 2$, we can determine a chain $\mathcal{D}_i$ of subsets of S (with $S \in \mathcal{D}_i$) and a function $\lambda_i : \mathcal{D}_i \to \mathbb{Z}$ such that

(47.45) $$w'_i = \sum_{W \in \mathcal{D}_i} \lambda_i(W)\chi^W$$

and such that $\lambda_i(W) > 0$ if $W \neq S$. We show that

(47.46) there exist $i \in \{1, 2\}$ and $W \in \mathcal{D}_i$ with $\lambda_i(W) > 2n$ and $\chi^W \notin L$.

Suppose not. Let $\mathcal{D}'_i := \{W \in \mathcal{D}_i \mid \chi^W \notin L\}$, and $\mathcal{D}''_i := \mathcal{D}_i \setminus \mathcal{D}'_i$, for $i = 1, 2$. So if $W \in \mathcal{D}'_i$, then $\lambda_i(W) \leq 2n$. This gives the contradiction:

(47.47) $$\begin{aligned} 4n^2 &\geq \|\sum_{i=1}^{2} \sum_{W \in \mathcal{D}'_i} \lambda_i(W)\chi^W\|_\infty = \|w' - \sum_{i=1}^{2} \sum_{W \in \mathcal{D}''_i} \lambda_i(W)\chi^W\|_\infty \\ &> \|\frac{5n^2}{\alpha}(w - y) - \sum_{i=1}^{2} \sum_{W \in \mathcal{D}''_i} \lambda_i(W)\chi^W\|_\infty - 1 \geq 5n^2 - 1. \end{aligned}$$

The last inequality holds as y minimizes $\|w - y\|_\infty$ over $y \in L$.

This shows (47.46). We can assume that $W \in \mathcal{D}'_1$ is such that $\lambda_1(W) > 2n$. Then:

(47.48) each optimum common base vector x of f_1 and f_2 satisfies $x(W) = f_1(W)$.

To see this, let

$$(47.49) \qquad \tilde{w} := \frac{5n^2}{\alpha}(w - y).$$

Replacing w by $\tilde{w}$ does not change the set of optimum common base vectors, since y belongs to L (implying (by our assumption (47.37)) that $y^{\mathsf{T}}x$ is the same for all common base vectors x of f_1 and f_2).

By Lemma 47.4α, there exists a good pair $\tilde{w}_1, \tilde{w}_2$ with $\tilde{w} = \tilde{w}_1 + \tilde{w}_2$ and

$$(47.50) \qquad \|\tilde{w}_i - w'_i\|_\infty \leq \|\tilde{w} - w'\|_1 < n$$

for $i = 1, 2$. Now for any $v \in W$ and $u \in S \setminus W$ we have $w'_1(v) > w'_1(u) + 2n$, as $\lambda_1(W) > 2n$, and as $\lambda_1(W') \geq 0$ for each $W' \in \mathcal{D}_1 \setminus \{S\}$. Hence, by (47.50), $\tilde{w}_1(v) > \tilde{w}_1(u)$. So (by the greedy method) any base vector x of f_1 maximizing $\tilde{w}_1^{\mathsf{T}}x$ satisfies $x(W) = f_1(W)$. This shows (47.48).

Let $\mathcal{C}_1 = \{U_1 \subset U_2 \subset \cdots \subset U_t = S\}$. For $j = 1, \ldots, t$, let $W_j := (W \cap U_j) \cup U_{j-1}$, where $U_0 := \emptyset$. Then $x(W_j) = f(W_j)$ for each optimum common base vector x (since W_j arises by taking unions and intersections from W, U_j, and U_{j-1}).

Moreover, $W_j \notin \mathcal{C}_1$ for at least one $j = 1, \ldots, t$, since

$$(47.51) \qquad \chi^W = \sum_{j=1}^{t}(\chi^{W_j} - \chi^{U_{j-1}})$$

while χ^W does not belong to L, implying that not all χ^{W_j} belong to L, and so some W_j does not belong to $\mathcal{C}_1$. So W_j can be added to $\mathcal{C}_1$, and we can iterate. ∎

From an optimum common base vector x, an optimum 'dual solution' w_1, w_2 can be derived, with a method of Cunningham and Frank [1985]. This gives:

Corollary 47.4a. *Problem* (47.25) *is solvable in strongly polynomial time.*

Proof. By Theorem 47.4, we can find a common base vector x of f_1 and f_2 maximizing $w^{\mathsf{T}}x$, in strongly polynomial time. Define a directed graph $D = (S, A)$ as follows.

For $i = 1, 2$, let A_i consist of all pairs (u, v) with $u, v \in S$ such that for each $U \subseteq V$:

(47.52) if $x(U) = f_i(U)$ and $u \in U$ then $v \in U$.

We can find A_i in strongly polynomial time, by finding the minimum of $f_i(U) - x(U)$ over subsets U of S with $u \in U$ and $v \notin U$ (with any strongly polynomial-time submodular function minimization algorithm).

Let D have arc set $A := A_1 \cup A_2^{-1}$ (taking two parallel arcs from u to v in case $(u, v) \in A_1$ and $(v, u) \in A_2$). Define a length function l on A by, for $(u, v) \in A$:

(47.53) $$l(u,v) := \begin{cases} w(v) - w(u) & \text{for } (u,v) \in A_1, \\ 0 & \text{for } (v,u) \in A_2. \end{cases}$$

We claim:

(47.54) D has no negative-length directed circuits.

For suppose that C is a negative-length directed circuit. We take such a C with $|AC|$ smallest. Then two consecutive arcs in C neither both belong to A_1 nor both belong to A_2^{-1}. For suppose that $a = (t,u)$ and $a' = (u,v)$ are in C and that they both belong to A_1. Then $(t,v) \in A_1$ and $l(a)+l(a') = l(t,v)$, contradicting the minimality of $|AC|$. This similarly gives a contradiction if $a, a' \in A_2^{-1}$.

So we can assume that C traverses the vertices $u_0, u_1, \ldots, u_k$ in this order, with $u_0 = u_k$, such that (u_{i-1}, u_i) belongs to A_1 if i is odd, and to A_2^{-1} if i is even. Let $X := \{u_1, u_3, \ldots, u_{k-1}\}$ and $Y := \{u_2, u_4, \ldots, u_k\}$. As C has negative length, we know $l(AC) < 0$, and hence $w(X) < w(Y)$.

By (47.52), for each $i = 1, 2$ and for each $U \subseteq V$ with $x(U) = f_i(U)$ we have $|U \cap Y| \geq |U \cap X|$. Hence there exists an $\varepsilon > 0$ such that the vector

(47.55) $$x' := x + \varepsilon(\chi^X - \chi^Y)$$

belongs to EP_{f_1} and to EP_{f_2}. So, since $x'(S) = x(S)$, x' is again a common base vector of f_1 and f_2. However, $w^\mathsf{T} x' = w^\mathsf{T} x + w(X) - w(Y) > w^\mathsf{T} x$, contradicting the fact that x maximizes $w^\mathsf{T} x$. This proves (47.54).

By Theorem 8.7, we can find a potential $p : S \to \mathbb{Z}$ for D with respect to l, in strongly polynomial time. Then p satisfies

(47.56) $$\begin{array}{ll} p(v) - p(u) \leq w(v) - w(u) & \text{for each } (u,v) \in A_1, \\ p(v) - p(u) \geq 0 & \text{for each } (u,v) \in A_2. \end{array}$$

Define $w_1 := w - p$ and $w_2 := p$. We show that w_1 and w_2 are as required in (47.25).

(47.56) implies that, for each $i = 1, 2$,

(47.57) if $(u,v) \in A_i$ then $w_i(v) \geq w_i(u)$.

We show that (47.57) implies that, for each $i = 1, 2$, x is a base vector of f_i maximizing $w_i^\mathsf{T} x$, as required. We may assume $i = 1$.

Let μ and ν be the minimum and maximum value (respectively) of the entries in w_1. For $j \in \mathbb{Z}$, let $U_j := \{v \in S \mid w_1(v) \geq \mu + j\}$. Then, taking $k := \nu - \mu$,

(47.58) $$w_1 = \mu \cdot \chi^S + \sum_{j=1}^{k} \chi^{U_j}.$$

Moreover,

(47.59) $x(U_j) = f_1(U_j)$ for each $j = 1, \ldots, k$.

Indeed, for all $s \in U_j$ and $t \in S \setminus U_j$ we have $(s,t) \notin A_1$ (by (47.57), since $w_1(t) < \mu + j \leq w_1(s)$). Hence, by definition of A_1, there is a set $T_{s,t}$ with $s \in T_{s,t}$, $t \notin T_{s,t}$, and $x(T_{s,t}) = f_1(T_{s,t})$. As the collection of sets U with $x(U) = f_1(U)$ is closed under taking unions and intersections, (47.59) follows.

Then for any base vector x' of f_1 we have

(47.60) $$w_1^{\mathsf{T}} x' = \mu x'(S) + \sum_{j=1}^{k} x'(U_j) \leq \mu f_1(S) + \sum_{j=1}^{k} f_1(U_j).$$

By (47.59), we here have equality throughout for $x' := x$, which proves that x maximizes $w_1^{\mathsf{T}} x$ over all base vectors of f_1. ∎

Theorem 47.4 implies for (nonextended) polymatroids (extending a result of Schönsleben [1980] for integer f_1 and f_2 for which there is a fixed K with $P_{f_1} \cap P_{f_2} \subseteq [0,K]^S$):

Corollary 47.4b. *Given submodular set functions f_1, f_2 on S (by value giving oracles) and a weight function $w \in \mathbb{Q}^S$, we can find a maximum-weight vector $x \in P_{f_1} \cap P_{f_2}$ in strongly polynomial time.*

Proof. We can assume that $f_1(\emptyset) = f_2(\emptyset) = 0$ and that f_1 and f_2 are nondecreasing (as we can replace $f_i(U)$ by $\min_{T \supseteq U} f_i(T)$). Extend S with a new element t to a set $S' := S + t$. Define set functions f_1' and f_2' on S' by:

(47.61) $$f_i'(U) := f_i(U) \text{ and } f_i'(U + t) := 0$$

for $U \subseteq S$ and $i = 1, 2$. Then f_1' and f_2' are submodular (using the nondecreasingness of f_1 and f_2). Moreover, consider any $x' \in \mathbb{R}^{S'}$ with $x'(S') = 0$. Let x be the restriction of x' to S. Then:

(47.62) $$x' \in EP_{f_i'} \text{ if and only if } x \in P_{f_i}.$$

Indeed, if $x' \in EP_{f_i'}$, then $x'(s) \geq 0$ for each $s \in S$, since $x'(S' - s) \leq f'(S' - s) = 0$ and $x'(S') = 0$, implying that $x(s) = x'(s) \geq 0$. Moreover, for each $U \subseteq S$ one has $x(U) = x'(U) \leq f_i'(U) = f_i(U)$. So $x \in P_{f_i}$.

Conversely, if $x \in P_f$, then for each $U \subseteq S$ one has $x'(U) = x(U) \leq f_i(U) = f_i'(U)$ and $x'(U + t) = x(U) - x(S) = -x(S \setminus U) \leq 0 = f_i'(U + t)$. So $x' \in EP_{f_i'}$. This proves (47.62).

Define $w' \in \mathbb{Q}^{S'}$ by $w'(v) := w(v)$ if $v \in S$, and $w'(t) := 0$. By Theorem 47.4, we can find a common base vector x' of f_1' and f_2' maximizing $w'^{\mathsf{T}} x'$ in strongly polynomial time. Let x be the restriction of x' to S. By (47.62), x maximizes $w^{\mathsf{T}} x$ over $P_{f_1} \cap P_{f_2}$. ∎

Similarly for maximum-weight common base vectors in (nonextended) polymatroids:

Corollary 47.4c. *Given submodular set functions f_1, f_2 on S (by value giving oracles) and a weight function $w \in \mathbb{Q}^S$, we can find a maximum-weight common base vector x of P_{f_1} and P_{f_2} in strongly polynomial time.*

Proof. Again, we can assume that $f_1(\emptyset) = f_2(\emptyset) = 0$ and that f_1 and f_2 are nondecreasing. Then the present corollary follows directly from Theorem 47.4, since

(47.63) $$P_{f_1} \cap P_{f_2} \cap \{x \mid x(S) = f_1(S)\} = EP_{f_1} \cap EP_{f_2} \cap \{x \mid x(S) = f_1(S)\}.$$

Indeed, if $x \in EP_{f_i}$ and $x(S) = f_i(S)$, then $x \geq \mathbf{0}$, since for any $s \in S$ one has $x_s = x(S) - x(S - s) \geq f_i(S) - f_i(S - s) \geq 0$. ∎

Back to extended polymatroids, Corollary 47.4b yields that we can optimize over the intersection of two extended polymatroids in strongly polynomial time:

Corollary 47.4d. *Given submodular set functions f_1, f_2 on S (by value giving oracles) and a weight function $w \in \mathbb{Q}^S_+$, we can find a maximum-weight vector $x \in EP_{f_1} \cap EP_{f_2}$ in strongly polynomial time.*

Proof. We may assume that $f_1(\emptyset) = f_2(\emptyset) = 0$. Let

(47.64) $$L := \max_{i=1,2}(|f_i(S)| + \sum_{s \in S} |f_i(\{s\})|).$$

Then $|f_i(U)| \leq L$ for each $i = 1, 2$ and $U \subseteq S$, since

(47.65) $$f_i(U) \leq \sum_{s \in U} f_i(\{s\}) \leq L$$

and

(47.66) $$f_i(U) \geq f_i(S) - f_i(S \setminus U) \geq f_i(S) - \sum_{s \in S \setminus U} f_i(\{s\}) \geq -L.$$

Let $K := |S| \cdot L + 1$. Then for any vertex x of $EP_{f_1} \cap EP_{f_2}$ and any $s \in S$:

(47.67) $$x(s) > -K,$$

since $x = A^{-1}b$ for some totally unimodular matrix A and some vector b whose entries are values of f_1 and f_2 (as in the proof of Theorem 46.1; observe that the entries of A^{-1} belong to $\{0, \pm 1\}$).

Define $f'_i(U) := f_i(U) + K \cdot |U|$. Then

(47.68) $$EP_{f'_i} = K \cdot \mathbf{1} + EP_{f_i}$$

for $i = 1, 2$. Hence, by (47.67), all vertices of $EP_{f'_1} \cap EP_{f'_2}$ are nonnegative. So any vector x maximizing $w^{\mathsf{T}}x$ over $P_{f'_1} \cap P_{f'_2}$ also maximizes $w^{\mathsf{T}}x$ over $EP_{f'_1} \cap EP_{f'_2}$. By Corollary 47.4b, x can be found in strongly polynomial time. ∎

47.5. Contrapolymatroids

Similar results hold for intersections of contrapolymatroids, by reduction to polymatroids. Given supermodular set functions g_1 and g_2 on S (by value giving oracles) and a weight function $w \in \mathbb{Q}^S$, we can find in strongly polynomial time:

(47.69) (i) a minimum-weight vector in $EQ_{g_1} \cap EQ_{g_2}$,
(ii) a minimum-weight common base vector of EQ_{g_1} and EQ_{g_2},
(iii) a minimum-weight vector in $Q_{g_1} \cap Q_{g_2}$, and
(iv) a minimum-weight common base vector of Q_{g_1} and Q_{g_2}.

Here (i) and (ii) follow from Corollary 47.4d and Theorem 47.4 applied to the submodular functions $-g_1$ and $-g_2$. Moreover, (iii) and (iv) follow by application of (i) and (ii) to the supermodular functions $\bar{g}_i$ given by $\bar{g}_i(U) = \max_{T \subseteq U} g_i(T)$ for $U \subseteq S$ and $i = 1, 2$ (assuming without loss of generality $g_1(\emptyset) = g_2(\emptyset) = 0$).

47.6. Intersecting a polymatroid and a contrapolymatroid

Let f be a submodular, and g a supermodular, set function on S. The results on polymatroid intersection also imply that

(47.70) a maximum-weight vector in the intersection of the extended polymatroid EP_f and the extended contrapolymatroid EQ_g can be found in strongly polynomial time,

assuming that we have value giving oracles for f and g.

To see this, we can assume that $f(\emptyset) = g(\emptyset) = 0$ and $g(S) \leq f(S)$. Let t be a new element. Define submodular set functions f_1 and f_2 on $S + t$ by:

(47.71) $f_1(U) := f(U)$, $f_1(U+t) := f(U) - g(S)$, $f_2(U) := f(S) - g(S \setminus U)$, $f_2(U + t) := -g(S \setminus U)$,

for $U \subseteq S$. Reset $f_1(S + t) := 0$. Then for each $x \in \mathbb{R}^S$ and $\lambda \in \mathbb{R}$:

(47.72) (x, λ) is a common base vector of EP_{f_1} and EP_{f_2}
$\iff \lambda = -x(S)$ and $x \in EP_f \cap EQ_g$.

To see necessity, let (x, λ) be a common base vector of EP_{f_1} and EP_{f_2}. As $f_1(S + t) = 0$, we have $\lambda = -x(S)$. Moreover, for any $U \subseteq S$, we have

(47.73) $x(U) \leq f_1(U) = f(U)$ and
$x(U) = x(S) - x(S \setminus U) = -\lambda - x(S \setminus U) \geq -f_2((S \setminus U) + t) = g(U)$.

So $x \in EP_f \cap EQ_g$.

To see sufficiency, assume $\lambda = -x(S)$ and $x \in EP_f \cap EQ_g$. Then for each $U \subseteq S$ we have:

(47.74)
$$\begin{aligned}&x(U) \le f(U) = f_1(U),\\ &x(U+t) = x(U) + \lambda = x(U) - x(S) \le f(U) - g(S) = f_1(U+t),\\ &x(U) = x(S) - x(S \setminus U) \le f(S) - g(S \setminus U) = f_2(U),\\ &x(U+t) = x(U) + \lambda = x(U) - x(S) = -x(S \setminus U) \le -g(S \setminus U)\\ &= f_2(U+t).\end{aligned}$$

So (x, λ) is a common base vector of EP_{f_1} and EP_{f_2}.

This shows (47.72), which implies that finding a minimum-weight vector in $EP_f \cap EQ_g$ amounts to finding a minimum-weight common base vector of EP_{f_1} and EP_{f_2}.

Similarly, we can find a modular function h satisfying $g \le h \le f$ in strongly polynomial time, if $g \le f$ (Frank's discrete sandwich theorem (Corollary 46.2b)). To see this, let f_1 and f_2 be as above, and find an (x, λ) in $EP_{f_1} \cap EP_{f_2}$ maximizing $x(S) + \lambda$. If $x(S) + \lambda \ge 0$, then $x \in EP_f \cap EQ_g$, that is x gives a modular function h with $g \le h \le f$.

47.6a. Further notes

Polymatroid intersection is a special case of submodular flow, as discussed in Chapter 60. We therefore refer for further algorithmic work to the notes in Section 60.3e.

A preflow-push algorithm for finding a maximum common vector in the intersection of two polymatroids was presented by Fujishige and Zhang [1992].

Tardos, Tovey, and Trick [1986] gave an improved version of Cunningham and Frank's polynomial-time algorithm for weighted polymatroid intersection. Fujishige [1978a] gave a (non-polynomial-time) algorithm for weighted polymatroid intersection. Optimizing over the intersection of a base polytope and an affine space was considered by Hartvigsen [1996,1998a,2001a].

Frank [1984c] and Fujishige and Iwata [2000] gave surveys.

Chapter 48

Dilworth truncation

If a submodular function f has $f(\emptyset) < 0$, the associated extended polymatroid is empty, as the conditions $x(U) \leq f(U)$ for all U include $x(\emptyset) < f(\emptyset)$. However, by ignoring the condition for $U = \emptyset$, the obtained polyhedron is yet an extended polymatroid, for a different submodular function, denoted by $\hat{f}$. This function $\hat{f}$ is called the *Dilworth truncation* of f.

48.1. If $f(\emptyset) < 0$

Let f be a submodular set function on S. If $f(\emptyset) < 0$, the associated extended polymatroid EP_f is empty. However, by ignoring the empty set, we yet obtain an extended polymatroid. (The interest in this goes back to Dilworth [1944].)

Consider the system

(48.1) $$x(U) \leq f(U) \text{ for } U \in \mathcal{P}(S) \setminus \{\emptyset\},$$

and the problem dual to maximizing $w^{\mathsf{T}}x$ over (48.1), for $w \in \mathbb{R}_+^S$:

(48.2) $$\min\{\sum_{U \in \mathcal{P}(S)\setminus\{\emptyset\}} y(U)f(U) \mid y \in \mathbb{R}_+^{\mathcal{P}(S)\setminus\{\emptyset\}}, \sum_{U \in \mathcal{P}(S)\setminus\{\emptyset\}} y(U)\chi^U = w\}.$$

Recall that a collection $\mathcal{F}$ of sets is called *laminar* if

(48.3) $$T \cap U = \emptyset \text{ or } T \subseteq U \text{ or } U \subseteq T \text{ for all } T, U \in \mathcal{F}.$$

Then a basic result of Edmonds [1970b] is:

Theorem 48.1. *If f is a submodular set function on S, then* (48.2) *has an optimum solution y with $\mathcal{F} := \{U \in \mathcal{P}(S) \setminus \{\emptyset\} \mid y(U) > 0\}$ laminar.*

Proof. Let $y : \mathcal{P}(S) \setminus \{\emptyset\} \to \mathbb{R}_+$ achieve this minimum, with

(48.4) $$\sum_{U \in \mathcal{P}(S)\setminus\{\emptyset\}} y(U)|U||S \setminus U|$$

as small as possible. Assume that $\mathcal{F}$ is not laminar, and choose $T, U \in \mathcal{F}$ violating (48.3). Let $\alpha := \min\{y(T), y(U)\}$. Decrease $y(T)$ and $y(U)$ by α, and increase $y(T \cap U)$ and $y(T \cup U)$ by α. Since

$$\chi^{T\cap U} + \chi^{T\cup U} = \chi^T + \chi^U, \tag{48.5}$$

y remains a feasible solution of (48.2). As moreover

$$f(T \cap U) + f(T \cup U) \leq f(T) + f(U), \tag{48.6}$$

f remains optimum. However, by Theorem 2.1, sum (48.4) decreases, contradicting our assumption. ∎

This implies that system (48.1) is TDI. More generally, it implies the box-total dual integrality of (48.1):

Corollary 48.1a. *For any submodular set function f on S, system* (48.1) *is box-totally dual integral.*

Proof. Consider some $w : S \to \mathbb{Z}_+$, and problem (48.2) dual to maximizing $w^\mathsf{T} x$ over (48.1). By Theorem 48.1, this minimum is attained by a y with $\mathcal{F} := \{U \in \mathcal{P}(S) \setminus \{\emptyset\} \mid y(U) > 0\}$ laminar. Hence the constraints corresponding to positive entries in y form a totally unimodular matrix (by Theorem 41.11). Therefore, by Theorem 5.35, (48.1) is box-TDI. ∎

Let EP'_f denote the solution set of (48.1). So EP'_f is nonempty for each submodular function f. As for integrality we have:

Corollary 48.1b. *If f is submodular and integer, then EP'_f is integer.*

Proof. Directly from Corollary 48.1a. ∎

In fact, as we shall see in Section 48.2, EP'_f is again an extended polymatroid.

48.2. Dilworth truncation

For each submodular function f, there exists a unique largest submodular function $\hat{f}$ with the property that $\hat{f}(U) \leq f(U)$ for each nonempty $U \subseteq S$, and $\hat{f}(\emptyset) = 0$. This follows from a method of Dilworth [1944].

Let f be a submodular set function on S. The *Dilworth truncation* $\hat{f} : \mathcal{P}(S) \to \mathbb{R}$ of f is given by:

$$\hat{f}(U) := \min\{\sum_{P\in\mathcal{P}} f(P) \mid \mathcal{P} \text{ is a partition of } U \text{ into nonempty sets}\} \tag{48.7}$$

for $U \subseteq S$. So $\hat{f}(\emptyset) = 0$ (as for $U = \emptyset$, only $\mathcal{P} = \emptyset$ qualifies in (48.7)). Dunstan [1976] showed:

Theorem 48.2. $\hat{f}$ *is submodular.*

Proof. Choose $T, U \subseteq S$, and let $\mathcal{P}$ and $\mathcal{Q}$ be partitions of T and U (respectively) into nonempty sets with

$$\hat{f}(T) = \sum_{P \in \mathcal{P}} f(P) \text{ and } \hat{f}(U) = \sum_{Q \in \mathcal{Q}} f(Q). \tag{48.8}$$

Consider the family $\mathcal{F}$ made by $\mathcal{P}$ and $\mathcal{Q}$ (taking a set twice if it occurs in both partitions). We can transform $\mathcal{F}$ iteratively into a laminar family, by replacing any $X, Y \in \mathcal{F}$ with $X \cap Y \neq \emptyset$ and $X \not\subseteq Y \not\subseteq X$ by $X \cap Y, X \cup Y$. In each iteration, the sum

$$\sum_{Z \in \mathcal{F}} f(Z) \tag{48.9}$$

does not increase (as f is submodular). As at each iteration the sum

$$\sum_{Z \in \mathcal{F}} |Z||S \setminus Z| \tag{48.10}$$

decreases (by Theorem 2.1), this process terminates. We end up with a laminar family $\mathcal{F}$.

The inclusionwise maximal sets in $\mathcal{F}$ form a partition $\mathcal{R}$ of $T \cup U$, and the remaining sets form a partition $\mathcal{S}$ of $T \cap U$. Therefore,

$$\begin{aligned} \hat{f}(T \cup U) + \hat{f}(T \cap U) &\leq \sum_{X \in \mathcal{R}} f(X) + \sum_{Y \in \mathcal{S}} f(Y) \\ &\leq \sum_{P \in \mathcal{P}} f(P) + \sum_{Q \in \mathcal{Q}} f(Q) = \hat{f}(T) + \hat{f}(U), \end{aligned} \tag{48.11}$$

showing that $\hat{f}$ is submodular. ∎

Lovász [1983c] observed that $\hat{f}$ is the unique largest among all submodular set functions g on S with $g(\emptyset) = 0$ and $g(U) \leq f(U)$ for $U \neq \emptyset$. Indeed, each subset U of S can be partitioned into nonempty sets $U_1, \ldots, U_t$ such that

$$g(U) \leq \sum_{i=1}^{t} g(U_i) \leq \sum_{i=1}^{t} f(U_i) = \hat{f}(U) \tag{48.12}$$

(the first inequality follows from the submodularity of g, as $g(\emptyset) = 0$).

Trivially, $EP_{\hat{f}} = EP'_f$. In particular, EP'_f is an extended polymatroid. Moreover, by (44.34),

Theorem 48.3.

$$\hat{f}(U) = \max\{x(U) \mid x \in EP'_f\}. \tag{48.13}$$

Proof. By (44.34), since $EP'_f = EP_{\hat{f}}$. ∎

$\hat{f}(U)$ can be computed in strongly polynomial time:

Theorem 48.4. *If a submodular set function f on S is given by a value giving oracle, then for each given $U \subseteq S$, $\hat{f}(U)$ can be computed in strongly polynomial time.*

Proof. We can assume that $U = S$. Order $S = \{s_1, \ldots, s_n\}$ arbitrarily. For $i = 1, \ldots, n$, define $U_i := \{s_1, \ldots, s_i\}$. Set $x := \mathbf{0}$. Iteratively, for $i = 1, \ldots, n$, determine

(48.14) $\mu := \min\{f(T) - x(T) \mid s_i \in T \subseteq U_i\}$

(with a submodular function minimization algorithm), and reset $x(s_i) := x(s_i) + \mu$.

We end up with $x \in EP'_f$ and for each $u \in S$ a subset T_u of S with $u \in T_u$ and $x(T_u) = f(T_u)$. As the collection of subsets T of S with $x(T) = f(T)$ is closed under unions and intersections of intersecting sets (cf. Theorem 44.2), we can modify the T_u in such a way that they form a partition $U_1, \ldots, U_k$ of S. Then $\hat{f}(S) = f(U_1) + \cdots + f(U_k)$, as x attains the maximum in (48.13). ∎

As a consequence, given a submodular set function f on S (by a value giving oracle), we can optimize over EP'_f in strongly polynomial time (by Corollary 44.3b, as $EP'_f = EP_{\hat{f}}$ and as we can compute $\hat{f}$).

48.2a. Applications and interpretations

Graphic matroids (Dilworth [1944], also Edmonds [1970b], Dunstan [1976]). Let $G = (V, E)$ be an undirected graph and let for each $F \subseteq E$, $f(F)$ be given by

(48.15) $f(F) := \left|\bigcup F\right| - 1.$

It is easily checked that the function f is submodular, and that the function $\hat{f}$ as given by (48.7) satisfies

(48.16) $\hat{f}(F) = |V|$ minus the number of components of the graph (V, F),

i.e., $\hat{f}$ is the rank function of the cycle matroid of G.

Geometric interpretation. The operation of making $\hat{f}$ from f can be interpreted geometrically as follows (Lovász [1977c], Mason [1977,1981]).

Let $\mathcal{F}$ be a collection of flats (subspaces) in a projective space, and define for each subset $\mathcal{F}'$ of $\mathcal{F}$, the rank $r(\mathcal{F}')$ by

(48.17) $r(\mathcal{F}') :=$ the (projective) dimension of $\bigcup \mathcal{F}'$.

One easily checks that r is nondecreasing and submodular and that $r(\emptyset) = 0$. Now let

(48.18) $f(\mathcal{F}') := r(\mathcal{F}') - 1$

for $\mathcal{F}' \subseteq \mathcal{F}$, and consider the function $\hat{f}$. Then $\hat{f}$ can be interpreted geometrically as follows. Let H be some hyperplane 'in general position' in the projective space. Then $\hat{f}(\mathcal{F}')$ is equal to the projective dimension of $H \cap \bigcup \mathcal{F}'$, i.e., $\hat{f}$ is as given by (48.17) if we replace $\mathcal{F}$ by $\{F \cap H \mid F \in \mathcal{F}\}$ (see Lovász [1977c] and Mason [1977, 1981]).

Rigidity. Let $M = (S, \mathcal{I})$ be a loopless matroid, with rank function r. Let d be a natural number. Define the set function f on S by

(48.19) $f(U) := d \cdot r(U) - d + 1,$

for $U \subseteq S$. Again, f is submodular and nondecreasing. Moreover, the function $\hat{f}$ is the rank function of a loopless matroid, as $\hat{f}(\{s\}) = f(\{s\}) = 1$ for all s in S.

Let $M_d = (S, \mathcal{I}_d)$ be this matroid. Since $EP'_f = EP_{\hat{f}}$, a subset I of S is independent in M_d if and only if

(48.20) $|U| \leq d \cdot r(U) - d + 1$

for all $U \subseteq I$.

In case M is the cycle matroid of a connected graph $G = (V, E)$, this relates to the following (cf. Crapo [1979] and Crapo and Whiteley [1978]). Let the vertices of G be placed 'in general position' in the d-dimensional Euclidean space. Make the edges 'rigid bars'. Suppose now that the whole graph G is rigid (which only depends on G and not on the embedding, since the vertices are 'in general position'). Then G is called *rigid (in d dimensions)*. It is not difficult to see that the minimal sets F of edges of G for which the subgraph (V, F) is rigid, form the bases of a matroid. For $d = 1$ this matroid is just the cycle matroid of G, as can be checked easily. Laman [1970] (cf. Asimow and Roth [1978,1979]) showed that for $d = 2$, a graph $G = (V, E)$ is a base (i.e., a minimal rigid graph), if and only if

(48.21) (i) $|E| = 2|V| - 3$,
(ii) $|E[U]| \leq 2|U| - 3$, for each $U \subseteq V$.

Now if M is the cycle matroid of a rigid graph G, with rank function r, then (48.21)(ii) is equivalent to

(48.22) $|F| \leq 2r(F) - 1$, for each subset F of E,

that is, by (48.20), to: E is independent in the matroid M_2, as given above. Condition (48.21)(i) implies that M_2 has rank $2r(E) - 1$. Hence, if G is rigid in 2 dimensions, then the bases of M_2 are the minimally rigid subgraphs of G in 2 dimensions.

In general, the matroid of rigid subgraphs of a graph $G = (V, E)$ (in d dimensions) has rank $d|V| - \binom{d+1}{2}$. However, it is not necessarily true that G is minimally rigid in d dimensions if and only if G has $d|V| - \binom{d+1}{2}$ edges and each subgraph (U, F) of G has at most $d|U| - \binom{d+1}{2}$ edges. For instance, if G arises from glueing two copies of the complete graph K_5 together in two vertices, and deleting the edge connecting these two vertices, then G is not rigid in 3 dimensions, but it satisfies the conditions given above for $d = 3$. (These conditions are easily seen to be necessary.)

More on the relation between rigidity and matroid union can be found in Whiteley [1988].

48.3. Intersection

Corollaries 48.1a and 48.1b on submodular functions f not necessarily satisfying $f(\emptyset) \geq 0$, can be extended to pairs of functions. Let f_1 and f_2 be submodular set functions on S, and consider the system

(48.23) $$x(U) \leq f_i(U) \text{ for } U \in \mathcal{P}(S) \setminus \{\emptyset\} \text{ and } i = 1, 2.$$

Then:

Theorem 48.5. *System* (48.23) *is box-totally dual integral.*

Proof. Choose $w \in \mathbb{Z}^S$, and consider the problem dual to maximizing $w^{\mathsf{T}} x$ over (48.23):

(48.24) $$\min\{\sum_{U \in \mathcal{P}(S)\setminus\{\emptyset\}} (y_1(U) f_1(U) + y_2(U) f_2(U)) \mid y_1, y_2 \in \mathbb{R}_+^{\mathcal{P}(S)\setminus\{\emptyset\}}, \sum_{U \in \mathcal{P}(S)\setminus\{\emptyset\}} (y_1(U) + y_2(U))\chi^U = w\}.$$

Let y_1, y_2 attain the minimum.

For $i \in \{1, 2\}$, define

(48.25) $$w_i := \sum_{U \in \mathcal{P}(S)\setminus\{\emptyset\}} y_i(U)\chi^U.$$

By Theorem 48.1, for each $i = 1, 2$,

(48.26) $$\min\{\sum_{U \in \mathcal{P}(S)\setminus\{\emptyset\}} y_i(U) f_i(U) \mid y_i \in \mathbb{R}_+^{\mathcal{P}(S)\setminus\{\emptyset\}}, \sum_{U \in \mathcal{P}(S)\setminus\{\emptyset\}} y_i(U)\chi^U = w_i\}$$

has an optimum solution y_i with $\mathcal{F}_i := \{U \mid y_i(U) > 0\}$ laminar.

These (modified) y_1, y_2 again are optimum in (48.24). As the constraints corresponding to positive components of y_1, y_2 give a totally unimodular matrix (by Theorem 41.11), Theorem 5.35 implies that system (48.23) is box-TDI. ∎

Theorem 48.5 implies primal integrality:

Corollary 48.5a. *If f_1 and f_2 are submodular and integer, then $EP'_{f_1} \cap EP'_{f_2}$ is box-integer.*

Proof. Directly from Theorem 48.5. ∎

Given submodular functions f_1 and f_2 (by value giving oracles), we can optimize over $EP'_{f_1} \cap EP'_{f_2}$ in strongly polynomial time (by Corollary 47.4d, as $EP'_{f_1} = EP_{\hat{f}_1}$ and $EP'_{f_2} = EP_{\hat{f}_2}$).

Chapter 49

Submodularity more generally

We now discuss a number of generalizations of submodular functions, namely those defined on a subcollection $\mathcal{C}$ of the collection of all subsets of a set S. The results are similar to those for submodular functions defined on all subsets on S. Often, the corresponding polyhedra form a polymatroid for some derived submodular function defined on all subsets of S.
We consider three kinds of collections, in order of increasing generality: lattice families, intersecting families, and crossing families.

49.1. Submodular functions on a lattice family

We first consider the generalization of submodular functions to those defined on a 'lattice family'.

Let S be a finite set. A family $\mathcal{C}$ of sets is called a *lattice family* if

(49.1) $\quad T \cap U, T \cup U \in \mathcal{C}$ for all $T, U \in \mathcal{C}$.

For a lattice family $\mathcal{C}$, a function $f : \mathcal{C} \to \mathbb{R}$ is called *submodular* if

(49.2) $\quad f(T \cap U) + f(T \cup U) \leq f(T) + f(U)$

for all $T, U \in \mathcal{C}$. Consider the system

(49.3) $\quad x(U) \leq f(U)$ for $U \in \mathcal{C}$,

and the problem dual to maximizing $w^{\mathsf{T}}x$ over (49.3), for $w \in \mathbb{R}^S$:

(49.4) $\quad \min\{\sum_{U \in \mathcal{C}} y(U)f(U) \mid y \in \mathbb{R}^{\mathcal{C}}_{+}, \sum_{U \in \mathcal{C}} y(U)\chi^U = w\}.$

Theorem 49.1. *Let $\mathcal{C}$ be a lattice family, $f : \mathcal{C} \to \mathbb{R}$ a submodular function, and $w \in \mathbb{R}^S$. Then (49.4) has an optimum solution y with $\mathcal{F} := \{U \in \mathcal{C} \mid y(U) > 0\}$ a chain.*

Proof. Let $y : \mathcal{C} \to \mathbb{R}_+$ achieve this minimum, with

(49.5) $\quad \sum_{U \in \mathcal{C}} y(U)|U||S \setminus U|$

as small as possible. Assume that $\mathcal{F}$ is not a chain, and choose $T, U \in \mathcal{F}$ with $T \not\subseteq U$ and $U \not\subseteq T$. Let $\alpha := \min\{y(T), y(U)\}$. Decrease $y(T)$ and $y(U)$ by α, and increase $y(T \cap U)$ and $y(T \cup U)$ by α. Since

$$\chi^{T\cap U} + \chi^{T\cup U} = \chi^T + \chi^U, \tag{49.6}$$

y remains a feasible solution of (49.4). As moreover

$$f(T \cap U) + f(T \cup U) \leq f(T) + f(U), \tag{49.7}$$

f remains optimum. However, by Theorem 2.1, sum (49.5) decreases, contradicting our assumption. ∎

This implies the box-total dual integrality of (49.3):

Corollary 49.1a. *If $\mathcal{C}$ is a lattice family and $f : \mathcal{C} \to \mathbb{R}$ is submodular, then system* (49.3) *is box-TDI.*

Proof. Consider some $w : \mathcal{C} \to \mathbb{Z}$, and problem (49.4) dual to maximizing $w^{\mathsf{T}}x$ over (49.3). By Theorem 49.1, this minimum is attained by a y with $\mathcal{F} := \{U \in \mathcal{C} \mid y(U) > 0\}$ a chain. So the constraints corresponding to positive components of y form a totally unimodular matrix (by Theorem 41.11). Hence by Theorem 5.35, (49.3) is box-TDI. ∎

For any $\mathcal{C} \subseteq \mathcal{P}(S)$ and $f : \mathcal{C} \to \mathbb{R}$, define:

$$\begin{aligned} P_f &:= \{x \in \mathbb{R}^S \mid x \geq \mathbf{0}, x(U) \leq f(U) \text{ for each } U \in \mathcal{C}\}, \\ EP_f &:= \{x \in \mathbb{R}^S \mid x(U) \leq f(U) \text{ for each } U \in \mathcal{C}\}. \end{aligned} \tag{49.8}$$

Then Corollary 49.1a implies:

Corollary 49.1b. *If $\mathcal{C}$ is a lattice family and $f : \mathcal{C} \to \mathbb{R}$ is submodular and integer, then EP_f is box-integer.*

Proof. Directly from Corollary 49.1a. ∎

Another consequence of Theorem 49.1 is that a submodular function f on a lattice family is uniquely determined by EP_f (given the lattice family):

Corollary 49.1c. *If $\mathcal{C}$ is a lattice family and $f : \mathcal{C} \to \mathbb{R}$ is submodular, then*

$$f(U) = \max\{x(U) \mid x \in EP_f\} \tag{49.9}$$

for each $U \in \mathcal{C}$.

Proof. Let $w := \chi^U$ and let y attain minimum (49.4), with $\mathcal{F} := \{T \in \mathcal{C} \mid y(T) > 0\}$ a chain. Since

$$\chi^U = w = \sum_{T \in \mathcal{F}} y(T)\chi^T, \tag{49.10}$$

we know that $\mathcal{F} = \{U\}$ and $y(U) = 1$. So the maximum in (49.9) is equal to $\sum_{T \in \mathcal{C}} y(T)f(T) = y(U)f(U) = f(U)$. ∎

We note that for any lattice family $\mathcal{C} \subseteq \mathcal{P}(S)$ with $\bigcup \mathcal{C} = S$, and any submodular function $f : \mathcal{C} \to \mathbb{R}$, the polytope P_f is a polymatroid. Indeed, define

(49.11) $$f'(U) := \min\{f(T) \mid T \in \mathcal{C}, T \supseteq U\}.$$

for $U \subseteq S$. Then f' is submodular, and $P_{f'} = P_f$.

49.2. Intersection

Also the intersection of two of the polyhedra EP_f is tractable. Let S be a finite set. For $i = 1, 2$, let $\mathcal{C}_i$ be a lattice family on S and let $f_i : \mathcal{C}_i \to \mathbb{R}$ be submodular. Consider the system

(49.12) $$\begin{aligned} &x(U) \leq f_1(U) \text{ for } U \in \mathcal{C}_1, \\ &x(U) \leq f_2(U) \text{ for } U \in \mathcal{C}_2. \end{aligned}$$

Then:

Corollary 49.1d. *System* (49.12) *is box-TDI.*

Proof. Choose $w \in \mathbb{R}^S$, and consider the problem dual to maximizing $w^{\mathsf{T}}x$ over (49.12):

(49.13) $$\begin{aligned} \min\{ &\sum_{U \in \mathcal{C}_1} y_1(U)f_1(U) + \sum_{U \in \mathcal{C}_2} y_2(U)f_2(U) \mid \\ &y_1 \in \mathbb{R}_+^{\mathcal{C}_1}, y_2 \in \mathbb{R}_+^{\mathcal{C}_2}, \sum_{U \in \mathcal{C}_1} y_1(U)\chi^U + \sum_{U \in \mathcal{C}_2} y_2(U)\chi^U = w\}. \end{aligned}$$

Let y_1, y_2 attain the minimum.

For $i \in \{1, 2\}$, define

(49.14) $$w_i := \sum_{U \in \mathcal{C}_i} y_i(U)\chi^U.$$

By Theorem 49.1, for each $i = 1, 2$,

(49.15) $$\min\{\sum_{U \in \mathcal{C}_i} y_i(U)f_i(U) \mid y_i \in \mathbb{R}_+^{\mathcal{C}_i}, \sum_{U \in \mathcal{C}_i} y_i(U)\chi^U = w_i\}$$

has an optimum solution y_i with $\mathcal{F}_i := \{U \in \mathcal{C}_i \mid y_i(U) > 0\}$ a chain.

These y_1, y_2 again are optimum in (49.13). So, by Theorem 41.11, the constraints corresponding to positive components of y form a totally unimodular matrix. Hence by Theorem 5.35, (49.12) is box-TDI. ∎

This implies primal integrality:

Corollary 49.1e. *If f_1 and f_2 are submodular and integer, then $EP_{f_1} \cap EP_{f_2}$ is box-integer.*

Proof. Directly from Corollary 49.1d. ∎

49.3. Complexity

To find the minimum of a submodular function f defined on a lattice family $\mathcal{C}$ in polynomial time, just an oracle telling if a set U belongs to $\mathcal{C}$, and if so, giving $f(U)$, is not sufficient: if $\mathcal{C} = \{\emptyset, T, S\}$ for some $T \subseteq S$, with $f(\emptyset) = f(S) = 0$ and $f(T) = -1$, we cannot find T by a polynomially bounded number of oracle calls. So we need to have more information on $\mathcal{C}$.

A lattice family $\mathcal{C}$ is fully characterized by the smallest set M and the largest set L in $\mathcal{C}$, together with the pre-order $\preceq$ on S defined by:

(49.16) $\quad u \preceq v \iff$ each $U \in \mathcal{C}$ containing v also contains u.

Then $\preceq$ is a pre-order (that is, it is reflexive and transitive). A subset U of S belongs to $\mathcal{C}$ if and only if $M \subseteq U \subseteq L$ and U is a lower ideal in $\preceq$ (that is, if $v \in U$ and $u \preceq v$, then $u \in U$).

Hence $\mathcal{C}$ has a description of size $O(|S|^2)$, such that for given $U \subseteq S$ one can test in polynomial time if U belongs to $\mathcal{C}$.

For $U \subseteq S$, define

(49.17) $\quad U^{\downarrow} := \{s \in S \mid \exists t \in U : s \preceq t\}$ and
$\quad U^{\uparrow} := \{s \in S \mid \exists t \in U : t \preceq s\}$.

Set

(49.18) $\quad v^{\uparrow} := \{v\}^{\uparrow},\ v^{\downarrow} := \{v\}^{\downarrow},\ \tilde{v} := v^{\uparrow} \cap v^{\downarrow}$.

For any $U \subseteq S$, let $\overline{U}$ be the (unique) smallest set in $\mathcal{C}$ containing $U \cap L$; that is,

(49.19) $\quad \overline{U} = (U \cap L)^{\downarrow} \cup M$.

So having L, M, and $\preceq$, the set $\overline{U}$ can be determined in polynomial time.

Determine a number $\alpha > 0$ such that

(49.20) $\quad \alpha \geq f(S \setminus v^{\uparrow}) - f((S \setminus v^{\uparrow}) \cup \tilde{v})$ and $\alpha \geq f(v^{\downarrow}) - f(v^{\downarrow} \setminus \tilde{v})$

for all $v \in L \setminus M$. Such an α can be found by at most $4|S|$ oracle calls.

Then α satisfies, for any $X, Y \in \mathcal{C}$ with $X \subseteq Y$:

(49.21) $\quad |f(Y) - f(X)| \leq \alpha |Y \setminus X|$.

To show this, we can assume that $Y \setminus X = \tilde{v}$ for some $v \in L \setminus M$. Then $f(Y) - f(X) \leq f(v^{\downarrow}) - f(v^{\downarrow} \setminus \tilde{v}) \leq \alpha$ and $f(Y) - f(X) \geq f((S \setminus v^{\uparrow}) \cup \tilde{v}) - f(S \setminus v^{\uparrow}) \geq -\alpha$, implying (49.21).

Now define a function $\bar{f} : \mathcal{P}(S) \to \mathbb{R}$ by:

(49.22) $\bar{f}(U) := f(\overline{U}) + \alpha|\overline{U} \triangle U|$

for $U \subseteq S$.

Then:

Theorem 49.2. *For any $\alpha > 0$ satisfying (49.20) for all $v \in L \setminus M$, the function $\bar{f}$ is submodular.*

Proof. First consider $T, U \subseteq L$. Then $T \subseteq \overline{T}$ and $U \subseteq \overline{U}$, and hence:

(49.23)
$$\begin{aligned}\bar{f}(T) + \bar{f}(U) &= f(\overline{T}) + \alpha|\overline{T} \setminus T| + f(\overline{U}) + \alpha|\overline{U} \setminus U| \\ &\geq f(\overline{T} \cap \overline{U}) + \alpha|(\overline{T} \cap \overline{U}) \setminus (T \cap U)| + f(\overline{T} \cup \overline{U}) + \alpha|(\overline{T} \cup \overline{U}) \setminus (T \cup U)| \\ &\geq f(\overline{T \cap U}) + \alpha|\overline{T \cap U} \setminus (T \cap U)| + f(\overline{T \cup U}) + \alpha|\overline{T \cup U} \setminus (T \cup U)| \\ &= \bar{f}(T \cap U) + \bar{f}(T \cup U).\end{aligned}$$

(The last inequality uses (49.21), since $\overline{T} \cap \overline{U} \supseteq \overline{T \cap U}$ (while $\overline{T} \cup \overline{U} = \overline{T \cup U}$).) Hence, for $T, U \subseteq S$ one has:

(49.24)
$$\begin{aligned}\bar{f}(T) + \bar{f}(U) &= \bar{f}(T \cap L) + \alpha|T \setminus L| + \bar{f}(U \cap L) + \alpha|U \setminus L| \\ &\geq \bar{f}((T \cap L) \cap (U \cap L)) + \bar{f}((T \cap L) \cup (U \cap L)) + \alpha|T \setminus L| + \alpha|U \setminus L| \\ &= \bar{f}((T \cap U) \cap L) + \bar{f}((T \cup U) \cap L) + \alpha|(T \cap U) \setminus L| + \alpha|(T \cup U) \setminus L| \\ &= \bar{f}(T \cap U) + \bar{f}(T \cup U).\end{aligned}$$

So $\bar{f}$ is submodular. ∎

The function $\bar{f}$ enables us to reduce optimization problems on submodular functions defined on a lattice family, to those defined on all subsets.

Minimization. By Theorem 45.1, the minimum of $\bar{f}$ can be found in strongly polynomial time. Hence

(49.25) if $\mathcal{C}$ is given by L, M, and $\preceq$, and a submodular function $f : \mathcal{C} \to \mathbb{R}$ is given by a value giving oracle, we can find a $U \in \mathcal{C}$ minimizing $f(U)$ in strongly polynomial time.

Indeed, if $\bar{f}$ attains its minimum at U, then $U \in \mathcal{C}$, since otherwise $\overline{U} \neq U$ and hence $\bar{f}(U) > \bar{f}(\overline{U})$ (as $\alpha > 0$), contradicting the fact that $\bar{f}$ attains its minimum at U. This shows (49.25).

Maximization over EP_f. Given a lattice family $\mathcal{C}$ of subsets of a set S, a submodular function $f : \mathcal{C} \to \mathbb{R}$, and a weight function $w \in \mathbb{Q}^S$, we can maximize $w^{\mathsf{T}}x$ over EP_f, by adapting the greedy algorithm as follows.

Note that $\max\{w^{\mathsf{T}}x \mid x \in EP_f\}$ is finite if and only if $w \geq \mathbf{0}$, $w(s) = 0$ for each $s \in S \setminus L$, and

(49.26) $u \preceq v$ implies $w(u) \geq w(v)$

for all $u, v \in S$. If (49.26) is not the case, the maximum value is infinite, since if $u \preceq v$, then for any $x \in EP_f$, the vector $x + \lambda(\chi^v - \chi^u)$ belongs to EP_f for all $\lambda \geq 0$. Now, if $w(v) > w(u)$, the weight increases to infinity along this line, and therefore the maximum value is ∞.

So we can check in strongly polynomial time if $\max\{w^{\mathsf{T}}x \mid x \in EP_f\}$ is finite, and therefore we can assume that it is finite. Moreover, we can assume that $L = S$, since $w(s) = 0$ for each $s \in S \setminus L$, and hence we can delete $S \setminus L$. Similarly, we can assume that $\emptyset \in \mathcal{C}$ and $f(\emptyset) = 0$. For if $\emptyset \in \mathcal{C}$ and $f(\emptyset) < 0$, then $EP_f = \emptyset$, and if $f(\emptyset) > 0$, we can reset $f(\emptyset) := 0$, without violating the submodularity and without modifying EP_f. If $\emptyset \notin \mathcal{C}$, then we can add $\emptyset$ to $\mathcal{C}$ and set $f(\emptyset) := 0$, again maintaining submodularity and EP_f. Finally, we can assume that $\preceq$ is a partial order, since if $u \preceq v \preceq u$, then by (49.26), $w(u) = w(v)$, and each set in $\mathcal{C}$ either contains both u and v, or neither of them. So we can merge u and v; and in fact we can merge any set $\tilde{v}$ to one element.

Now let $\leq$ be a linear order such that for any u, v, if $u \preceq v$ or $w(u) > w(v)$, then $u \leq v$. By (49.26), the latter defines a partial order. So $\leq$ is a linear extension of it, and hence can be found in strongly polynomial time.

Let $S = \{s_1, \ldots, s_n\}$ with $s_1 < s_2 < \cdots < s_n$. For $i = 0, \ldots, n$, define $U_i := \{s_1, \ldots, s_i\}$. As $\leq$ is a linear extension of $\preceq$, each U_i is a lower ideal of $\preceq$, and hence each U_i belongs to $\mathcal{C}$. Define $x(s_i) := f(U_i) - f(U_{i-1})$ for $i = 1, \ldots, n$. Then x maximizes $w^{\mathsf{T}}x$ over EP_f.

To see this, let $\bar{f}$ be defined as above. Then by Theorem 44.3, x belongs to $EP_{\bar{f}}$ (as f and $\bar{f}$ coincide on each U_i), and hence x belongs to EP_f. To see that x is optimum, we have for any $z \in EP_f$:

$$\begin{aligned}(49.27)\qquad w^{\mathsf{T}}z &= \sum_{i=1}^{n-1} z(U_i)(w(s_i) - w(s_{i+1})) + z(S)w(s_n)\\ &\leq \sum_{i=1}^{n-1} f(U_i)(w(s_i) - w(s_{i+1})) + f(S)w(s_n)\\ &= \sum_{i=1}^{n} w(s_i)(f(U_i) - f(U_{i-1})) = \sum_{i=1}^{n} w(s_i)x(s_i) = w^{\mathsf{T}}x.\end{aligned}$$

This also gives a dual solution to the corresponding LP-formulation of the problem.

Maximization over intersections. Let $\mathcal{C}_1$ and $\mathcal{C}_2$ be lattice families of subsets of S and let f_1 and f_2 be submodular functions on $\mathcal{C}_1$ and $\mathcal{C}_2$ respectively. Let $\mathcal{C}_i$ be specified by L_i, M_i, and $\preceq_i$.

Find a number $\alpha > 0$ satisfying (49.20) for both $f = f_1$ and $f = f_2$. So by (49.21), $\alpha|S| + \max_{i=1,2}|f_i(L_i)|$ is an upper bound on $|f_i(U)|$ for each $i \in \{1, 2\}$ and each $U \in \mathcal{C}_i$. Define

$$(49.28)\qquad K := |S|(\alpha|S| + \max_{i=1,2}|f_i(L_i)|).$$

Now for $i = 1, 2$ and $U \subseteq S$, let $\bar{f}_i(U) := f_i(\overline{U}) + K|\overline{U} \triangle U|$ (where $\overline{U}$ is taken with respect to $\mathcal{C}_i$). So $\bar{f}_1$ and $\bar{f}_2$ are submodular (by Theorem 49.2). Then:

$$(49.29)\qquad \max\{w^{\mathsf{T}}x \mid x \in EP_{f_1} \cap EP_{f_2}\} = \max\{w^{\mathsf{T}}x \mid x \in EP_{\bar{f}_1} \cap EP_{\bar{f}_2}\},$$

if the first maximum is finite. Clearly $\geq$ holds in (49.29), since $EP_{\bar{f}_i} \subseteq EP_{f_i}$ for $i = 1, 2$. To see equality, each face of $EP_{f_1} \cap EP_{f_2}$ is determined by equations $x(U) = f_i(U)$ for $i = 1, 2$ and $U \in \mathcal{D}_i$, where $\mathcal{D}_i$ is a chain of sets in $\mathcal{C}_i$. So it is determined by a system of linear equations with totally unimodular constraint set and right-hand sides determined by function values of f_1 and f_2. So each face contains a vector x with $|x_s| \leq K$ for all $s \in S$ (by (49.28), since the inverse of a nonsingular totally unimodular matrix has all its entries in $\{0, \pm 1\}$). But any such x belongs to $EP_{\bar{f}_1} \cap EP_{\bar{f}_2}$, since for $i = 1, 2$ and $U \subseteq S$, we have:

(49.30) $$x(U) \leq x(\overline{U}) + K|\overline{U} \triangle U| \leq f_i(\overline{U}) + K|\overline{U} \triangle U| = \bar{f}_i(U)$$

(where $\overline{U}$ is taken with respect to $\mathcal{C}_i$). So we have (49.29).

Therefore, by Corollary 47.4d, we can maximize $w^{\mathsf{T}}x$ over $EP_{f_1} \cap EP_{f_2}$ in strongly polynomial time. Note that for any $w \in \mathbb{Q}^S$, we can decide in strongly polynomial time if the first maximum in (49.29) is finite. For this, we should decide if there exist $w_1, w_2 \in \mathbb{Q}^S_+$ such that $w = w_1 + w_2$ and such that for $i = 1, 2$: $w_i(s) = 0$ for $s \in S \setminus L_i$ and $u \preceq_i v$ implies $w_i(u) \geq w_i(v)$ for all u, v. This can be reduced to checking if a certain digraph with lengths has no negative-length directed circuit.

49.4. Submodular functions on an intersecting family

We next consider functions defined on a broader class of collections, the intersecting families, where the function satisfies a restricted form of submodularity. It yields an extension of the *Dilworth truncation* studied in Chapter 48.

A family $\mathcal{C}$ of sets is called an *intersecting family* if for all $T, U \in \mathcal{C}$ one has:

(49.31) if $T \cap U \neq \emptyset$, then $T \cap U, T \cup U \in \mathcal{C}$.

Let $\mathcal{C}$ be an intersecting family. A function $f : \mathcal{C} \to \mathbb{R}$ is called *submodular on intersecting pairs*, or *intersecting submodular*, if

(49.32) $$f(T) + f(U) \geq f(T \cap U) + f(T \cup U)$$

for all $T, U \in \mathcal{C}$ with $T \cap U \neq \emptyset$.

Consider the system

(49.33) $x(U) \leq f(U)$ for $U \in \mathcal{C}$,

and the problem dual to maximizing $w^{\mathsf{T}}x$ over (49.33), for $w \in \mathbb{R}^S$:

(49.34) $$\min\{\sum_{U \in \mathcal{C}} y(U) f(U) \mid y \in \mathbb{R}^{\mathcal{C}}_+, \sum_{U \in \mathcal{C}} y(U) \chi^U = w\}.$$

Recall that a collection $\mathcal{F}$ of sets is called *laminar* if

(49.35) $T \cap U = \emptyset$ or $T \subseteq U$ or $U \subseteq T$, for all $T, U \in \mathcal{F}$.

A basic result (proved with a method due to Edmonds [1970b]) is:

Theorem 49.3. *Let $\mathcal{C}$ be an intersecting family of subsets of a set S, let $f : \mathcal{C} \to \mathbb{R}$ be intersecting submodular and let $w \in \mathbb{R}^S$. Then* (49.34) *has an optimum solution y with $\mathcal{F} := \{U \in \mathcal{C} \mid y(U) > 0\}$ laminar.*

Proof. Let $y : \mathcal{C} \to \mathbb{R}_+$ achieve this minimum, with

(49.36) $$\sum_{U \in \mathcal{C}} y(U)|U||S \setminus U|$$

as small as possible. Assume that $\mathcal{F}$ is not laminar, and choose $T, U \in \mathcal{F}$ violating (49.35). Let $\alpha := \min\{y(T), y(U)\}$. Decrease $y(T)$ and $y(U)$ by α, and increase $y(T \cap U)$ and $y(T \cup U)$ by α. Since

(49.37) $$\chi^{T \cap U} + \chi^{T \cup U} = \chi^T + \chi^U,$$

y remains a feasible solution of (49.34). As moreover

(49.38) $$f(T \cap U) + f(T \cup U) \leq f(T) + f(U),$$

f remains optimum. However, by Theorem 2.1, sum (49.36) decreases, contradicting our assumption. ∎

It gives the box-total dual integrality of (49.33):

Corollary 49.3a. *Let $\mathcal{C}$ be an intersecting family of subsets of a set S and let $f : \mathcal{C} \to \mathbb{R}$ be intersecting submodular. Then system* (49.33) *is box-TDI.*

Proof. Consider problem (49.34) dual to maximizing $w^{\mathsf{T}}x$ over (49.33). By Theorem 49.3, this minimum is attained by a y with $\mathcal{F} := \{U \in \mathcal{C} \mid y(U) > 0\}$ laminar. As the matrix of constraints corresponding to $\mathcal{F}$ is totally unimodular (Theorem 41.11), Theorem 5.35 gives the corollary. ∎

49.5. Intersection

Again, these results can be extended in a natural way to pairs of functions, by derivation from Theorem 49.3.

Corollary 49.3b. *For $i = 1, 2$, let $\mathcal{C}_i$ be an intersecting family of subsets of a set S and let $f_i : \mathcal{C}_i \to \mathbb{R}$ be intersecting submodular. Then the system*

(49.39) $$\begin{aligned} &x(U) \leq f_1(U) \textit{ for } U \in \mathcal{C}_1, \\ &x(U) \leq f_2(U) \textit{ for } U \in \mathcal{C}_2, \end{aligned}$$

is box-TDI.

Proof. Choose $w \in \mathbb{R}^S$, and consider the problem dual to maximizing $w^{\mathsf{T}}x$ over (49.39):

$$(49.40) \qquad \min\{\sum_{U\in\mathcal{C}_1} y_1(U)f_1(U) + \sum_{U\in\mathcal{C}_2} y_2(U)f_2(U) \mid y_1 \in \mathbb{R}_+^{\mathcal{C}_1}, y_2 \in \mathbb{R}_+^{\mathcal{C}_2}, \sum_{U\in\mathcal{C}_1} y_1(U)\chi^U + \sum_{U\in\mathcal{C}_2} y_2(U)\chi^U = w\}.$$

Let y_1, y_2 attain the minimum. For $i \in \{1,2\}$, define

$$(49.41) \qquad w_i := \sum_{U\in\mathcal{C}_i} y_i(U)\chi^U.$$

By Theorem 49.3,

$$(49.42) \qquad \min\{\sum_{U\in\mathcal{C}_i} y_i(U)f_i(U) \mid y_i \in \mathbb{R}_+^{\mathcal{C}_i}, \sum_{U\in\mathcal{C}_i} y_i(U)\chi^U = w_i\}$$

has an optimum solution y_i with $\mathcal{F}_i := \{U \in \mathcal{C}_i \mid y_i(U) > 0\}$ laminar.

As $\mathcal{F}_1 \cup \mathcal{F}_2$ determine a totally unimodular matrix (by Theorem 41.11), Theorem 5.35 implies that system (49.39) is box-TDI. ∎

This implies the integrality of polyhedra:

Corollary 49.3c. *If f_1 and f_2 are integer, then $EP_{f_1} \cap EP_{f_2}$ is box-integer.*

Proof. Directly from Corollary 49.3b. ∎

49.6. From an intersecting family to a lattice family

Let $\mathcal{C}$ be an intersecting family of subsets of a set S and let $f : \mathcal{C} \to \mathbb{R}$ be submodular on intersecting pairs. Let $\check{\mathcal{C}}$ be the collection of all unions of sets in $\mathcal{C}$. Since $\mathcal{C}$ is closed under unions of intersecting sets, $\check{\mathcal{C}}$ is equal to the collection of *disjoint* unions of nonempty sets in $\mathcal{C}$. It is not difficult to show that $\check{\mathcal{C}}$ is a lattice family and that $\emptyset \in \check{\mathcal{C}}$.

Call a partition *proper* if its classes are nonempty. Define $\check{f} : \check{\mathcal{C}} \to \mathbb{R}$ by:

$$(49.43) \qquad \check{f}(U) := \min\{\sum_{P\in\mathcal{P}} f(P) \mid \mathcal{P} \subseteq \mathcal{C} \text{ is a proper partition of } U\}.$$

for $U \in \check{\mathcal{C}}$. So $\check{f}(\emptyset) = 0$. Then (Dunstan [1976]):

Theorem 49.4. *$\check{f}$ is submodular.*

Proof. Choose $T, U \in \check{\mathcal{C}}$, and let $\mathcal{P}$ and $\mathcal{Q}$ be partitions of T and U (respectively) into nonempty sets in $\mathcal{C}$ with

$$(49.44) \qquad \check{f}(T) = \sum_{P \in \mathcal{P}} f(P) \text{ and } \check{f}(U) = \sum_{Q \in \mathcal{Q}} f(Q).$$

Consider the family $\mathcal{F}$ made by the union of $\mathcal{P}$ and $\mathcal{Q}$ (taking a set twice if it occurs in both partitions). We can transform $\mathcal{F}$ iteratively into a laminar family, by replacing any $X, Y \in \mathcal{F}$ with $X \cap Y \neq \emptyset$ and $X \not\subseteq Y \not\subseteq X$ by $X \cap Y, X \cup Y$. In each iteration, the sum

$$(49.45) \qquad \sum_{Z \in \mathcal{F}} f(Z)$$

does not increase (as f is submodular on intersecting pairs). As at each iteration the sum

$$(49.46) \qquad \sum_{Z \in \mathcal{F}} |Z||S \setminus Z|$$

decreases (by Theorem 2.1), this process terminates. We end up with a laminar family $\mathcal{F}$.

The inclusionwise maximal sets in $\mathcal{F}$ form a partition $\mathcal{R}$ of $T \cup U$, and the remaining sets form a partition $\mathcal{S}$ of $T \cap U$. Therefore,

$$(49.47) \qquad \check{f}(T \cup U) + \check{f}(T \cap U) \leq \sum_{X \in \mathcal{R}} f(X) + \sum_{Y \in \mathcal{S}} f(Y) \leq \sum_{P \in \mathcal{P}} f(P) + \sum_{Q \in \mathcal{Q}} f(Q) = \check{f}(T) + \check{f}(U),$$

showing that $\check{f}$ is submodular. ∎

Trivially, if $\emptyset \notin \mathcal{C}$ or if $f(\emptyset) \geq 0$, then $EP_{\check{f}} = EP_f$. Hence, by (49.9),

$$(49.48) \qquad \check{f}(U) = \max\{x(U) \mid U \in EP_f\}.$$

As we shall see in Section 49.7, this enables us to calculate $\check{f}$ from a value giving oracle, using the greedy algorithm.

49.7. Complexity

The results of the previous section enable us to reduce algorithmic problems on intersecting submodular functions, to those on submodular functions on lattice families.

If $\mathcal{C}$ is an intersecting family on S, then for each $s \in S$, the collection $\mathcal{C}_s := \{U \in \mathcal{C} \mid s \in U\}$ is a lattice family. So (like in Section 49.3) we can assume that $\mathcal{C}$ is given by a representation of $\mathcal{C}_s$ for each $s \in S$, in terms of the pre-order $\preceq_s$ given by: $u \preceq_s v$ if and only if each set in $\mathcal{C}$ containing s and v also contains u, and by $M_s := \bigcap \mathcal{C}_s$ and $L_s := \bigcup \mathcal{C}_s$; next to that we should know if $\emptyset$ belongs to $\mathcal{C}$.

We can derive the information on $\check{\mathcal{C}}$ as follows:

(49.49) $\quad \bigcap \check{\mathcal{C}} = \emptyset$, $\bigcup \check{\mathcal{C}} = \bigcup_{s \in S} L_s$; $u \preceq v$ if and only if $u \in M_v$.

So we can decide in polynomial time if a set U belongs to $\check{\mathcal{C}}$.

For any intersecting submodular function f on $\mathcal{C}$, the restriction f_s of f to $\mathcal{C}_s$ is submodular. So by the results of Section 49.3, we can find a set minimizing f in strongly polynomial time.

For any $U \in \check{\mathcal{C}}$, we can calculate $\check{f}(U)$, as defined in (49.43), in strongly polynomial time, having a value giving oracle for f. To see this, we use (49.48). We can assume that $\emptyset \notin \mathcal{C}$.

Order the elements of U as $t_1, \ldots, t_k$ such that if $L_{t_j} \subset L_{t_i}$, then $j < i$. For $i = 0, \ldots, k$, let $U_i := L_{t_1} \cup \cdots \cup L_{t_i}$. So $U_k = U$.

Initially, set $x(t) := 0$ for each $t \in U$. Next, for $i = 1, \ldots, k$, calculate

(49.50) $\quad \mu := \min\{f(T) - x(T) \mid T \in \mathcal{C}, t_i \in T \subseteq U_i\}$,

and reset $x(t_i) := x(t_i) + \mu$. We prove, by induction on i, that for $i = 0, 1, \ldots, k$ we have, after processing $t_1, \ldots, t_i$:

(49.51)
(i) $x(T) \leq f(T)$ for each $T \in \mathcal{C}$ with $T \subseteq U_i$,
(ii) for each $j = 1, \ldots, i$ there exists a $T \in \mathcal{C}$ with $t_j \in T \subseteq U_i$ and $x(T) = f(T)$.

For $i = 0$ this is trivial. Let $i \geq 1$. Consider any $T \in \mathcal{C}$ with $T \subseteq U_i$. If $t_i \in T$, then $x(T) \leq f(T)$, as at processing t_i we have added μ to $x(t_i)$. If $t_i \notin T$, then $T \subseteq U_{i-1}$. For suppose that there exists a $t_j \in T$ with $t_j \notin U_{i-1}$. So $j > i$ and $t_j \in L_{t_i}$, and therefore $L_{t_j} \subseteq L_{t_i}$, implying $L_{t_j} = L_{t_i}$ (since if $L_{t_j} \subset L_{t_i}$, then $j < i$). But then $t_i \in L_{t_j} \subseteq T$, contradicting the fact that $t_i \notin T$. So $T \subseteq U_{i-1}$. As $t_i \notin T$, $x(T)$ did not change at processing t_i, and hence we know $x(T) \leq f(T)$ by induction. This proves (49.51)(i).

To see (49.51)(ii), choose $j \leq i$. If $j = i$, there exists after processing t_i a T as required, as we have added μ to $x(t_i)$. If $j < i$, by induction there exists a $T \in \mathcal{C}$ with $t_j \in T \subseteq U_{i-1}$ and $x(T) = f(T)$ before processing t_i. If $t_i \notin T$, $x(T) = f(T)$ is maintained at processing t_i. If $t_i \in T$, then $t_i \in U_{i-1}$, and so $t_i \in L_{t_j}$ for some $j < i$. Hence $L_{t_i} \subseteq L_{t_j}$. So, by the choice of the order of U, $L_{t_i} = L_{t_j}$. Hence before processing t_i we have $x(T') \leq f(T')$ for each $T' \subseteq U_i$. So, as $x(T) = f(T)$ and $t_i \in T$, $x(t_i)$ is not modified at processing t_i. Therefore, $x(T) = f(T)$ holds also after processing t_i. This proves (49.51)(ii).

This shows (49.51), which gives, taking $i = k$, that $x(T) \leq f(T)$ for each $T \in \mathcal{C}$ with $T \subseteq U$, and that for each $t \in U$, we have a T containing t with $x(T) = f(T)$. We can replace any two T and T' with $T \cap T' \neq \emptyset$ by $T \cup T'$. We end up with a partition $\mathcal{T}$ of U with $x(U) = \sum_{T \in \mathcal{T}} f(T)$. Hence we know

(49.52) $\quad \check{f}(U) \geq x(U) = \sum_{T \in \mathcal{T}} x(T) = \sum_{T \in \mathcal{T}} f(T) \geq \check{f}(U)$,

and therefore we have equality throughout.

Having this, we can reduce the problem of maximizing $w^{\mathsf{T}} x$ over EP_f, where f is intersecting submodular, to that of maximizing $w^{\mathsf{T}} x$ over $EP_{\check{f}}$, which can be done in strongly polynomial time by the results of Section 49.3. Similarly for intersections of two such polyhedra EP_{f_1} and EP_{f_2}.

49.8. Intersecting a polymatroid and a contrapolymatroid

For an intersecting family $\mathcal{C}$, a function $g : \mathcal{C} \to \mathbb{R}$ is called *supermodular on intersecting pairs*, or *intersecting supermodular*, if $-g$ is intersecting submodular.

Let S be a finite set. Let $\mathcal{C}$ and $\mathcal{D}$ be collections of subsets of S and let $f : \mathcal{C} \to \mathbb{R}$ and $g : \mathcal{D} \to \mathbb{R}$. Consider the system

(49.53)
$$\begin{aligned} x(U) &\le f(U) \quad \text{for } U \in \mathcal{C},\\ x(U) &\ge g(U) \quad \text{for } U \in \mathcal{D}. \end{aligned}$$

Theorem 49.5. *If $\mathcal{C}$ and $\mathcal{D}$ are intersecting families, $f : \mathcal{C} \to \mathbb{R}$ is intersecting submodular, and $g : \mathcal{D} \to \mathbb{R}$ is intersecting supermodular, then system* (49.53) *is box-TDI.*

Proof. Choose $w \in \mathbb{Z}^S$, and consider the dual problem of maximizing $w^{\mathsf{T}}x$ over (49.53):

(49.54)
$$\begin{aligned} \min\{\sum_{U\in\mathcal{C}} y(U)f(U) - \sum_{U\in\mathcal{D}} z(U)g(U) \mid \\ y \in \mathbb{R}_+^{\mathcal{C}}, z \in \mathbb{R}_+^{\mathcal{D}}, \sum_{U\in\mathcal{C}} y(U)\chi^U - \sum_{U\in\mathcal{D}} z(U)\chi^U = w\}. \end{aligned}$$

Let y, z attain this minimum. Define

(49.55)
$$u := \sum_{U\in\mathcal{C}} y(U)\chi^U \text{ and } v := \sum_{U\in\mathcal{D}} z(U)\chi^U.$$

So y attains

(49.56)
$$\min\{\sum_{U\in\mathcal{C}} y(U)f(U) \mid y \in \mathbb{R}_+^{\mathcal{C}}, \sum_{U\in\mathcal{C}} y(U)\chi^U = u\}$$

and z attains

(49.57)
$$\max\{\sum_{U\in\mathcal{D}} z(U)g(U) \mid z \in \mathbb{R}_+^{\mathcal{D}}, \sum_{U\in\mathcal{D}} z(U)\chi^U = v\}.$$

By Theorem 49.3, (49.56) has an optimum solution y with $\mathcal{F} := \{U \in \mathcal{C} \mid y(U) > 0\}$ laminar. Similarly, (49.57) has an optimum solution z with $\mathcal{G} := \{U \in \mathcal{D} \mid z(U) > 0\}$ laminar. Now $\mathcal{F}$ and $\mathcal{G}$ determine a totally unimodular submatrix (by Theorem 41.11), and hence by Theorem 5.35, (49.53) is box-TDI. ∎

49.9. Submodular functions on a crossing family

Finally, we consider submodular functions defined on a crossing family. The results discussed above for submodular functions on intersecting families do not all transfer to crossing families. But certain restricted versions still hold.

A family $\mathcal{C}$ of subsets of a set S is called a *crossing family* if for all $T, U \in \mathcal{C}$ one has:

(49.58) if $T \cap U \neq \emptyset$ and $T \cup U \neq S$, then $T \cap U, T \cup U \in \mathcal{C}$.

A function $f : \mathcal{C} \to \mathbb{R}$, defined on a crossing family $\mathcal{C}$, is called *submodular on crossing pairs*, or *crossing submodular*, if for all $T, U \in \mathcal{C}$ with $T \cap U \neq \emptyset$ and $T \cup U \neq S$:

(49.59) $f(T) + f(U) \geq f(T \cap U) + f(T \cup U)$.

In general, the system

(49.60) $x(U) \leq f(U)$ for $U \in \mathcal{C}$

is *not* TDI. For instance, if $S = \{1, 2, 3\}$, $\mathcal{C} = \{\{1, 2\}, \{1, 3\}, \{2, 3\}\}$, from S, and $f(U) := 1$ for each $U \in \mathcal{C}$, then (49.60) not even determines an integer polyhedron (as $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})^{\mathsf{T}}$ is a vertex of it).

However, for any $k \in \mathbb{R}$, the system

(49.61) $x(U) \leq f(U)$ for $U \in \mathcal{C}$,
$x(S) = k$

is box-TDI. This can be done by reduction to Corollary 49.3a. Similarly for pairs of such functions. This was shown by Frank [1982b,1984a] and Fujishige [1984e].

Let $\mathcal{C}$ be a crossing family of subsets of a set S. Let $\hat{\mathcal{C}}$ be the collection of all nonempty intersections of sets in $\mathcal{C}$ (we allow the intersection of 0 sets, so $S \in \hat{\mathcal{C}}$). Since $\mathcal{C}$ is a crossing family, we know

(49.62) $\hat{\mathcal{C}} = \{U \mid U \neq \emptyset; \exists \mathcal{P} \subseteq \mathcal{C} : \mathcal{P} \text{ is a copartition of } S \setminus U\}$,

where a *copartition* of U is a collection $\mathcal{P}$ of subsets of S such that the collection $\{S \setminus T \mid T \in \mathcal{P}\}$ is a partition of U. We call the copartition *proper* if $T \neq S$ for each $T \in \mathcal{P}$.

Note that, in (49.62), restricting $\mathcal{P}$ to proper copartitions of $S \setminus U$, does not modify $\hat{\mathcal{C}}$. We allow $\mathcal{P} = \emptyset$, so $S \in \hat{\mathcal{C}}$.

Define $\hat{f} : \hat{\mathcal{C}} \to \mathbb{R}$ by:

(49.63) $\hat{f}(U) := \min\{\sum_{P \in \mathcal{P}} f(P) \mid \mathcal{P} \subseteq \mathcal{C} \text{ is a proper copartition of } S \setminus U\}$

for $U \in \hat{\mathcal{C}}$. So $\hat{f}(S) = 0$. Then:

Theorem 49.6. *$\hat{\mathcal{C}}$ is an intersecting family and $\hat{f}$ is submodular on intersecting pairs.*

Proof. Define, for $s \in S$,

(49.64) $\qquad \hat{\mathcal{C}}_s := \{U \in \hat{\mathcal{C}} \mid s \in U\}$ and $\mathcal{D} := \{S \setminus U \mid s \in U \in \mathcal{C}\}$.

As $\mathcal{C}$ is crossing, $\mathcal{D}$ is intersecting. Hence $\check{\mathcal{D}}$ is a lattice family. As $\hat{\mathcal{C}}_s = \{S \setminus U \mid U \in \check{\mathcal{D}}\}$, also $\hat{\mathcal{C}}_s$ is a lattice family. As this is true for each $s \in S$, $\hat{\mathcal{C}}$ is intersecting.

To prove that $\hat{f}$ is intersecting submodular, it suffices to show that for each $s \in S$, the restriction $\hat{f}_s$ of $\hat{f}$ to $\hat{\mathcal{C}}_s$ is submodular. Define $g : \mathcal{D} \to \mathbb{R}$ by $g(U) := f(S \setminus U)$ for $U \in \mathcal{D}$. Then g is intersecting submodular (as f is crossing submodular). Hence, by Theorem 49.4, $\check{g}$ is submodular on $\check{\mathcal{D}}$. As $\hat{f}_s(U) = \check{g}(S \setminus U)$ for $U \in \hat{\mathcal{C}}_s$, $\hat{f}_s$ is submodular. ∎

Fujishige [1984e] showed that the following box-TDI result can be derived from Corollary 49.3a:

Theorem 49.7. *Let $\mathcal{C}$ be a crossing family of subsets of S, let $f : \mathcal{C} \to \mathbb{R}$ be crossing submodular, and let $k \in \mathbb{R}$. Then system* (49.61) *is box-TDI and determines the polyhedron of maximal vectors of $EP_{f'}$ for some submodular function f' defined on a lattice family.*

Proof. We can assume that $k = 0$, since, choosing any $t \in S$ and resetting $f(U) := f(U) - k$ for all $U \in \mathcal{C}$ with $t \in U$, does not change the box-total dual integrality of (49.61). We can also assume that $\emptyset \notin \mathcal{C}$.

The box-total dual integrality of (49.61) follows from that of

(49.65) $$\begin{aligned} &x(U) \le \hat{f}(U) \quad \text{for } U \in \hat{\mathcal{C}},\\ &x(S) = 0, \end{aligned}$$

as (49.65) and (49.61) have the same solution set, and as each constraint in (49.65) is a nonnegative integer combination of constraints in (49.61). The box-total dual integrality of (49.65) follows from Corollary 49.3a (using Theorem 5.25). It also shows that the solution set of (49.61) is the set of maximal vectors of $EP_{f'}$ for some submodular function f' defined on a lattice family. ∎

Frank and Tardos [1984b] observed that this implies a relation with matroids:

Corollary 49.7a. *If $\mathcal{C}$ is a crossing family of subsets of a set S, $f : \mathcal{C} \to \mathbb{Z}$ is crossing submodular, and $k \in \mathbb{Z}$, then the collection*

(49.66) $\qquad \{B \subseteq S \mid |B| = k, |B \cap U| \le f(U) \text{ for each } U \in \mathcal{C}\}$

forms the collection of bases of a matroid (if nonempty).

Proof. Directly from Theorem 49.7, using the observations on (44.43) and (49.11). ∎

Similarly, the box-total dual integrality of pairs of such systems follows:

Theorem 49.8. *For $i = 1, 2$, let $\mathcal{C}_i$ be a crossing family of subsets of a set S, let $f_i : \mathcal{C}_i \to \mathbb{R}$ be crossing submodular, and let $k \in \mathbb{R}$. Then the system*

$$\begin{aligned} \text{(49.67)} \qquad & x(U) \leq f_1(U) \text{ for } U \in \mathcal{C}_1, \\ & x(U) \leq f_2(U) \text{ for } U \in \mathcal{C}_2, \\ & x(S) = k, \end{aligned}$$

is box-TDI.

Proof. Similar to the proof of the previous theorem, by reduction to Corollary 49.3b. ∎

This implies the integrality of polyhedra:

Corollary 49.8a. *For $i = 1, 2$, let $\mathcal{C}_i$ be a crossing family of subsets of a set S, let $f_i : \mathcal{C}_i \to \mathbb{R}$ be crossing submodular, and let $k \in \mathbb{R}$. If f_1, f_2, and k are integer, system* (49.67) *determines a box-integer polyhedron.*

Proof. Directly from Theorem 49.8. ∎

49.10. Complexity

The reduction given in the proof of Theorem 49.6 also enables us to calculate $\hat{f}(U)$ from a value giving oracle for f, similar to the proof in Section 49.7. We assume that $\mathcal{C}$ is given by descriptions of the lattice families $\mathcal{C}_{s,t} := \{U \in \mathcal{C} \mid s \in U, t \notin U\}$ as in Section 49.3.

Note that

$$\text{(49.68)} \qquad EP_{\hat{f}} \cap \{x \mid x(S) = 0\} = EP_f \cap \{x \mid x(S) = 0\}.$$

This follows from the fact that if $x \in EP_f$ and $x(S) = 0$, then for any $U \in \hat{\mathcal{C}}$ and any proper copartition $\mathcal{P} = \{U_1, \ldots, U_p\}$ of $S \setminus U$ with $f(U) = \sum_{P \in \mathcal{P}} f(P)$, one has:

$$\begin{aligned} \text{(49.69)} \qquad x(U) &= -x(S \setminus U) = -\sum_{P \in \mathcal{P}} x(S \setminus P) = \sum_{P \in \mathcal{P}} x(P) \leq \sum_{P \in \mathcal{P}} f(P) \\ &= f(U). \end{aligned}$$

Having this, the problem of maximizing $w^{\mathsf{T}}x$ over $EP_f \cap \{x \mid x(S) = 0\}$, where f is crossing submodular, is reduced to the problem of maximizing $w^{\mathsf{T}}x$ over $EP_{\hat{f}} \cap \{x \mid x(S) = 0\}$. The latter problem can be solved in strongly polynomial time by the results of Section 49.7. Similar results hold for intersections of two such polyhedra:

Theorem 49.9. *For crossing families $\mathcal{C}_1, \mathcal{C}_2$ of subsets of a set S, crossing submodular functions $f_1 : \mathcal{C}_1 \to \mathbb{Q}$ and $f_2 : \mathcal{C}_2 \to \mathbb{Q}$, $k \in \mathbb{Q}$, and $w \in \mathbb{Q}^S$, one* *can find an $x \in EP_{f_1} \cap EP_{f_2}$ with $x(S) = k$ and maximizing $w^{\mathsf{T}}x$ in strongly polynomial time.*

Proof. From the above. ∎

If $\mathcal{C}$ is a crossing family and $f : \mathcal{C} \to \mathbb{Q}$ is crossing submodular, then we can find its minimum value in polynomial time, as, for each $s, t \in S$, we can minimize f over the lattice family $\{U \in \mathcal{C} \mid s \in U, t \notin U\}$, and take the minimum of all these minima, and of the values in $\emptyset$ and S (if in $\mathcal{C}$).

Hence we can decide in polynomial time if a given vector $x \in \mathbb{Q}^S$ belongs to EP_f, by testing if the minimum value of the crossing submodular function $g : \mathcal{C} \to \mathbb{Q}$ defined by

$$g(U) := f(U) - x(U) \tag{49.70}$$

for $U \in \mathcal{C}$, is nonnegative.

49.10a. Nonemptiness of the base polyhedron

Let $\mathcal{C}$ be a crossing family of subsets of a set S and let $f : \mathcal{C} \to \mathbb{R}$ be crossing submodular. We give a theorem of Fujishige [1984e] characterizing when EP_f contains a vector x with $x(S) = 0$. If $S \in \mathcal{C}$ and $f(S) = 0$, then the set $EP_f \cap \{x \mid x(S) = 0\}$ is called the *base polyhedron* of f.

To give the characterization, again call a collection $\mathcal{P} \subseteq \mathcal{C}$ a *copartition* of S if the collection $\{S \setminus U \mid U \in \mathcal{P}\}$ is a partition of S.

Theorem 49.10. *EP_f contains a vector x satisfying $x(S) = 0$ if and only if*

$$\sum_{U \in \mathcal{P}} f(U) \geq 0 \tag{49.71}$$

for each partition or copartition $\mathcal{P} \subseteq \mathcal{C}$ of S. If moreover f is integer, there exists an integer such vector x.

Proof. The condition is necessary, since if $x \in EP_f$ satisfies $x(S) = 0$ and $\mathcal{P} \subseteq \mathcal{C}$ is a partition of S, then

$$\sum_{U \in \mathcal{P}} f(U) \geq \sum_{U \in \mathcal{P}} x(U) = x(S) = 0. \tag{49.72}$$

Similarly, if $\mathcal{P} \subseteq \mathcal{C}$ is a copartition of S, then

$$\sum_{U \in \mathcal{P}} f(U) \geq \sum_{U \in \mathcal{P}} x(U) = \sum_{U \in \mathcal{P}} (x(S) - x(S \setminus U)) = |\mathcal{P}|x(S) - x(S) = 0. \tag{49.73}$$

To see sufficiency, let $\mathcal{D} := \hat{\mathcal{C}}$ and $g := \hat{f}$ (cf. (49.62)). By Theorem 49.6, $\mathcal{D}$ is an intersecting family and g is intersecting submodular. Moreover, $S \in \mathcal{D}$. Let $\mathcal{E} := \check{\mathcal{D}}$ and $h := \check{g}$. By Theorem 49.4, $\mathcal{E}$ is a lattice family and h is submodular.

Now if EP_h contains a vector x with $x(S) = 0$, then $x \in EP_g$, and hence $x \in EP_f$ (using $x(S) = 0$). So it suffices to show that EP_h contains a vector x with $x(S) = 0$.

By Corollary 49.1c, in order to show this, it suffices to show that $h(S) \geq 0$. The solution can be taken integer if f (hence h) is integer.

Suppose $h(S) < 0$. Since $h = \check{g}$, $\mathcal{D}$ contains a proper partition $\mathcal{P}$ of S with

$$h(S) = \sum_{U \in \mathcal{P}} g(U). \tag{49.74}$$

Since $g = \hat{f}$, for each $U \in \mathcal{P}$, $\mathcal{C}$ contains a proper copartition $\mathcal{Q}_U$ of $S \setminus U$ such that

$$g(U) = \sum_{T \in \mathcal{Q}_U} f(T). \tag{49.75}$$

Let $\mathcal{F}$ be the family consisting of the union of the $\mathcal{Q}_U$ over $U \in \mathcal{P}$, taking multiplicities into account. Then

(49.76) (i) all elements of S are contained in the same number of sets in $\mathcal{F}$;
(ii) $\sum_{T \in \mathcal{F}} f(T) < 0$.

Now apply the following operation as often as possible to $\mathcal{F}$: if $T, W \in \mathcal{F}$ with $T \cap W \neq \emptyset$, $T \cup W \neq S$, and $T \not\subseteq W \not\subseteq T$, replace T and W by $T \cap W$ and $T \cup W$. This maintains (49.76) and decreases $\sum_{T \in \mathcal{F}} |T||S \setminus T|$ (by Theorem 2.1). So the process terminates, and we end up with a *cross-free family*: for all $T, W \in \mathcal{F}$ we have $T \subseteq W$ or $W \subseteq T$ or $T \cap W = \emptyset$ or $T \cup W = S$.

We show that $\mathcal{F}$ contains a partition or copartition $\mathcal{P}$ of S. By (49.71), $\mathcal{F} \setminus \mathcal{P}$ again satisfies (49.76), and hence we can repeat. We end up with $\mathcal{F}$ empty, a contradiction.

To show that $\mathcal{F}$ contains a partition or copartition of S, choose $U \in \mathcal{F}$. If $U = \emptyset$ or $U = S$ we are done (taking $\mathcal{P} := \{U\}$). So we can assume that $\emptyset \neq U \neq S$. Let $\mathcal{X}$ be the collection of inclusionwise maximal subsets of $S \setminus U$ that belong to $\mathcal{F}$. Let $\mathcal{Y}$ be the collection of inclusionwise minimal supersets of $S \setminus U$ that belong to $\mathcal{F}$. Since $\mathcal{F}$ is cross-free and $U \neq \emptyset$, the sets in $\mathcal{X}$ are pairwise disjoint. Similarly, the complements of the sets in $\mathcal{Y}$ are pairwise disjoint.

If $\bigcup \mathcal{X} = S \setminus U$, then $\mathcal{X} \cup \{U\}$ is a partition of S as required. If $\bigcap \mathcal{Y} = S \setminus U$, then $\mathcal{Y} \cup \{U\}$ is a copartition of S as required. So we can assume that there exist $s \in (S \setminus U) \setminus \bigcup \mathcal{X}$ and $t \in U \cap \bigcap \mathcal{Y}$. Since each element of S is contained in the same number of sets in $\mathcal{F}$, and since $s \notin U$, and $t \in U$, there exists a $T \in \mathcal{F}$ with $s \in T$ and $t \notin T$. So $T \not\subseteq U \not\subseteq T$.

Hence $T \cap U = \emptyset$ or $T \cup U = S$. However, if $T \cap U = \emptyset$, then T is contained in some set in $\mathcal{X}$, and hence $s \in T \subseteq \bigcup \mathcal{X}$, a contradiction. If $T \cup U = S$, then T contains some set in $\mathcal{Y}$, and hence $t \notin T \supseteq \bigcap \mathcal{Y}$, again a contradiction. ∎

This theorem will be used in proving Theorem 61.8.

Fujishige and Tomizawa [1983] characterized the vertices of the base polyhedron of a submodular function defined on a lattice family.

49.11. Further results and notes

49.11a. Minimizing a submodular function over a subcollection of a lattice family

In Section 45.7 we saw that the minimum of a submodular function over the odd subsets can be found in strongly polynomial time. A generalization of minimizing a submodular function over the odd subsets (cf. Section 45.7), was given by Grötschel, Lovász, and Schrijver [1981,1984a] (the latter paper corrects a serious flaw in the first paper found by A. Frank). This was extended by Goemans and Ramakrishnan [1995] to the following.

Let $\mathcal{C}$ be a lattice family and let $\mathcal{D}$ be a subcollection of $\mathcal{C}$ with the following property:

(49.77) for all $X, Y \in \mathcal{C} \setminus \mathcal{D}$: $X \cap Y \in \mathcal{D} \iff X \cup Y \in \mathcal{D}$.

Examples are: $\mathcal{D} := \{X \in \mathcal{C} \mid |X| \not\equiv q (\text{mod } p)\}$ for some natural numbers p, q, and $\mathcal{D} := \mathcal{C} \setminus \mathcal{A}$ for some antichain or some sublattice $\mathcal{A} \subseteq \mathcal{C}$.

To prove that for a submodular function on $\mathcal{C}$, the minimum over $\mathcal{D}$ can be found in strongly polynomial time, Goemans and Ramakrishnan gave the following interesting lemma:

Lemma 49.11α. *Let $\mathcal{C}$ be a lattice family, let f be a submodular function on $\mathcal{C}$, let $\mathcal{D} \subseteq \mathcal{C}$ satisfy* (49.77)*, and let U minimize $f(U)$ over $U \in \mathcal{D}$. If $U \neq \emptyset$, then there exists a $u \in U$ such that $f(W) \geq f(U)$ for each subset W of U with $W \in \mathcal{C}$ and $u \in W$.*

Proof. Suppose not. Then for each $u \in U$ there exists a $W_u \in \mathcal{C}$ satisfying $u \in W_u \subseteq U$ and $f(W_u) < f(U)$. Choose each W_u inclusionwise maximal with this property. Then

(49.78) $$f(\bigcap_{u \in T} W_u) < f(U)$$

for each nonempty $T \subseteq U$. To prove this, choose a counterexample T with $|T|$ minimal. Then $|T| > 1$, since $f(W_u) < f(U)$ for each $u \in U$. Choose $t \in T$. Since $\bigcap_{u \in T} W_u \neq \bigcap_{u \in T - t} W_u$ by the minimality of T, we know that $\bigcap_{u \in T-t} W_u \not\subseteq W_t$, and hence W_t is a proper subset of $(\bigcap_{u \in T-t} W_u) \cup W_t$. So by the maximality of W_t, $f((\bigcap_{u \in T-t} W_u) \cup W_t) \geq f(U)$. Hence

(49.79) $$\begin{aligned} f(U) &\leq f(\bigcap_{u \in T} W_u) = f((\bigcap_{u \in T-t} W_u) \cap W_t) \\ &\leq f(\bigcap_{u \in T-t} W_u) + f(W_t) - f((\bigcap_{u \in T-t} W_u) \cup W_t) \\ &< f(U) + f(U) - f(U) = f(U), \end{aligned}$$

a contradiction.

This shows (49.78), which implies

(49.80) $$\bigcap_{u \in T} W_u \notin \mathcal{D}$$

for each nonempty $T \subseteq U$.

This can be extended to:

(49.81) $$X := (\bigcap_{u \in T} W_u) \cap (\bigcup_{u \in V} W_u) \notin \mathcal{D}$$

for all disjoint $T, V \subseteq U$ with V nonempty. Suppose to the contrary that $X \in \mathcal{D}$. Choose such X with $|V|$ minimal. By (49.80), $|V| \geq 2$. Choose $v \in V$. The minimality of V gives

(49.82) $$(\bigcap_{u\in T} W_u)\cap W_v \notin \mathcal{D} \text{ and } (\bigcap_{u\in T} W_u)\cap(\bigcup_{u\in V\setminus\{v\}} W_u)\notin \mathcal{D}.$$

By assumption, the union of these sets belongs to $\mathcal{D}$. Hence, by (49.77), also their intersection belongs to $\mathcal{D}$; that is

(49.83) $$(\bigcap_{u\in T\cup\{v\}} W_u)\cap(\bigcup_{u\in V\setminus\{v\}} W_u)\in \mathcal{D}$$

This contradicts the minimality of $|V|$.

This proves (49.81), which gives for $T := \emptyset$ and $V := U$ a contradiction, since then $X = U \in \mathcal{D}$. ∎

This lemma is used in proving the following theorem, where we assume that $\mathcal{C}$ is given as in Section 49.3, f is given by a value giving oracle, and $\mathcal{D}$ is given by an oracle telling if any given set in $\mathcal{C}$ belongs to $\mathcal{D}$:

Theorem 49.11. *Given a submodular function f on a lattice family $\mathcal{C}$, and a subcollection $\mathcal{D}$ of $\mathcal{C}$ satisfying* (49.77)*, a set U minimizing $f(U)$ over $U \in \mathcal{D}$ can be found in strongly polynomial time.*

Proof. We describe the algorithm. For all distinct $s,t \in S$ define

(49.84) $$\mathcal{C}_{s,t} := \{U \in \mathcal{C} \mid s \in U, t \notin U\}.$$

Let $U_{s,t}$ be the inclusionwise minimal set minimizing f over $\mathcal{C}_{s,t}$. ($U_{s,t}$ can be found by minimizing a slight perturbation of f.) Choose in

(49.85) $$\{\emptyset, S\} \cup \{U_{s,t} \mid s,t \in S\}$$

a $U \in \mathcal{D}$ minimizing f. Then U minimizes f over $\mathcal{D}$.

To see this, we must show that set (49.85) contains a set minimizing f over $\mathcal{D}$. Let W be a set minimizing f over $\mathcal{D}$, with $|W|$ minimal. If $W \in \{\emptyset, S\}$ we are done. So we can assume that $W \notin \{\emptyset, S\}$. By Lemma 49.11$\alpha$ (applied to the function $\tilde{f}(X) := f(S\setminus X)$ for $X \subseteq S$), there exists an element $t \in S\setminus W$ such that each $T \supseteq W$ with $t \notin T$ satisfies $f(T) \geq f(W)$.

The lemma also gives the existence of an $s \in W$ such that each $T \subset W$ with $s \in T$ satisfies $f(T) > f(W)$. Indeed, for small enough $\varepsilon > 0$, W minimizes $f(X)+\varepsilon|X|$ over $X \in \mathcal{D}$. Hence, by Lemma 49.11α, there exists an $s \in W$ such that each $T \subseteq W$ with $s \in T$ satisfies $f(T)+\varepsilon|T| \geq f(W)+\varepsilon|W|$. This implies $f(T) > f(W)$ if $T \neq W$.

We show that $W = U_{s,t}$. Indeed, W minimizes f over $\mathcal{C}_{s,t}$, since

(49.86) $$f(U_{s,t}) \geq f(W\cap U_{s,t}) + f(W\cup U_{s,t}) - f(W) \geq f(W)+f(W)-f(W) = f(W).$$

Moreover, $W \subseteq U_{s,t}$, as otherwise $W\cap U_{s,t} \subset W$, implying that the second inequality in (49.86) would be strict.

So $f(W) = f(U_{s,t})$, and hence, by the minimality of $U_{s,t}$, we have $W = U_{s,t}$. ∎

It is interesting to note that this algorithm implies that the set $\{\emptyset, S\} \cup \{U_{s,t} \mid s,t \in S\}$ contains a set minimizing f over $\mathcal{D}$, for any nonempty subcollection $\mathcal{D}$ of $\mathcal{C}$ satisfying (49.77).

Goemans and Ramakrishnan showed that if $\mathcal{C}$ and $\mathcal{D}$ are *symmetric* (that is, $U \in \mathcal{C} \iff S \setminus U \in \mathcal{C}$, and similarly for $\mathcal{D}$) and $\emptyset \notin \mathcal{D}$, then (49.77) is equivalent to: if $X, Y \in \mathcal{C} \setminus \mathcal{D}$ are disjoint, then $X \cup Y \in \mathcal{C} \setminus \mathcal{D}$.

Related work was reported by Benczúr and Fülöp [2000].

49.11b. Generalized polymatroids

We now describe a generalization, given by Frank [1984b], that comprises sub- and supermodular functions, and (extended) polymatroids and contrapolymatroids. (Hassin [1978,1982] described the case $\mathcal{C} = \mathcal{D} = \mathcal{P}(S)$.)

Let $\mathcal{C}$ and $\mathcal{D}$ be intersecting families of subsets of a finite set S and let $f : \mathcal{C} \to \mathbb{R}$ and $g : \mathcal{D} \to \mathbb{R}$. We say that the pair (f, g) is *paramodular* if

(49.87)
(i) f is submodular on intersecting pairs,
(ii) g is supermodular on intersecting pairs,
(iii) if $T \in \mathcal{C}$ and $U \in \mathcal{D}$ with $T \setminus U \neq \emptyset$ and $U \setminus T \neq \emptyset$, then $T \setminus U \in \mathcal{C}$ and $U \setminus T \in \mathcal{D}$, and

$$f(T \setminus U) - g(U \setminus T) \leq f(T) - g(U).$$

If (f, g) is paramodular, the solution set P of the system (for $x \in \mathbb{R}^S$):

$$\text{(49.88)} \qquad \begin{aligned} x(U) &\leq f(U) && \text{for } U \in \mathcal{C}, \\ x(U) &\geq g(U) && \text{for } U \in \mathcal{D}, \end{aligned}$$

is called a *generalized polymatroid* (*determined by* (f, g)).

Generalized polymatroids generalize polymatroids (where $g(U) = 0$ for each $U \subseteq S$), extended polymatroids (where $\mathcal{D} = \emptyset$), contrapolymatroids (where $\mathcal{C} = \emptyset$ and $g(\{s\}) \geq 0$ for each $s \in S$), and extended contrapolymatroids (where $\mathcal{C} = \emptyset$).

The intersection of a generalized polymatroid with a 'box' $\{x \mid d \leq x \leq c\}$ (for $d, c \in \mathbb{R}^S$) is again a generalized polymatroid: we can add $\{s\}$ to $\mathcal{C}$ and to $\mathcal{D}$ if necessary, and (re)define $f(\{s\}) := w(s)$ and $g(\{s\}) := d(s)$, if necessary. This transformation does not violate the paramodularity of (f, g).

Another transformation is as follows. Let $P \subseteq \mathbb{R}^S$ be a generalized polymatroid and let $\kappa, \lambda \in \mathbb{R}$. Let t be a new element and let $S' := S \cup \{t\}$. Let P' be the polyhedron in $\mathbb{R}^{S'}$ given by

$$\text{(49.89)} \qquad P' := \{(x, \eta) \mid x \in P, \lambda \leq x(S) + \eta \leq \kappa\}.$$

Then P' again is a generalized polymatroid, determined by the functions obtained by extending $\mathcal{C}$ and $\mathcal{D}$ with S' and extending f, g with the values $f(S') := \kappa$ and $g(S') := \lambda$.

The class of generalized polymatroids is closed under projections. That is, for any generalized polymatroid $P \subseteq \mathbb{R}^S$ and any $t \in S$, the set

$$\text{(49.90)} \qquad P' := \{x \in \mathbb{R}^{S-t} \mid \exists \eta : (x, \eta) \in P\}$$

is again a generalized polymatroid. This will be shown as Corollary 49.13c.

The following theorem will imply that system (49.88) is TDI. Hence, if f and g are integer, then P is integer.

Theorem 49.12. *System* (49.88) *is box-TDI.*

Proof. Let t be a new element. Define

$$(49.91) \qquad \mathcal{B} := \mathcal{C} \cup \{(S \setminus D) \cup \{t\} \mid D \in \mathcal{D}\}.$$

Then $\mathcal{B}$ is a crossing family of subsets of $S \cup \{t\}$. Define $e : \mathcal{B} \to \mathbb{R}$ by: $e(C) := f(C)$ for $C \in \mathcal{C}$ and $e((S \setminus D) \cup \{t\}) := -g(D)$ for $D \in \mathcal{D}$. Then e is crossing submodular. Hence, by Theorem 49.7, system

$$(49.92) \qquad \begin{aligned} &x(U) \leq e(U) && \text{for } U \in \mathcal{B}, \\ &x(S \cup \{t\}) = 0, \end{aligned}$$

is box-TDI. Therefore, by Theorem 5.27, system (49.88) is box-TDI. ∎

This gives for the integrality of generalized polymatroids:

Corollary 49.12a. *If (f, g) is paramodular and f and g are integer, the generalized polymatroid is box-integer.*

Proof. Directly from Theorem 49.12. ∎

More generally one has the box-total dual integrality of the system

$$(49.93) \qquad \begin{aligned} &x(U) \leq f_1(U) && \text{for } U \in \mathcal{C}_1, \\ &x(U) \geq g_1(U) && \text{for } U \in \mathcal{D}_1, \\ &x(U) \leq f_2(U) && \text{for } U \in \mathcal{C}_2, \\ &x(U) \geq g_2(U) && \text{for } U \in \mathcal{D}_2, \end{aligned}$$

for pairs of paramodular pairs (f_i, g_i):

Corollary 49.12b. *For $i = 1, 2$, let $\mathcal{C}_i$ and $\mathcal{D}_i$ be intersecting families and let $f_i : \mathcal{C}_i \to \mathbb{R}$, $g_i : \mathcal{D}_i \to \mathbb{R}$ form a paramodular pair. Then system* (49.93) *is box-TDI.*

Proof. Similar to the proof of Theorem 49.12, by reduction to Theorem 49.8. ∎

This gives for primal integrality:

Corollary 49.12c. *If f_1, g_1, f_2 and g_2 are integer, the intersection of the associated generalized polymatroids is box-integer.*

Proof. Directly from Corollary 49.12b. ∎

Another consequence is the following box-TDI result of McDiarmid [1978]:

Corollary 49.12d. *Let f_1 and f_2 be submodular set functions on a set S and let $\lambda, \kappa \in \mathbb{R}$. Then the system*

$$(49.94) \qquad \begin{aligned} &x(U) \leq f_1(U) && \text{for } U \subseteq S, \\ &x(U) \leq f_2(U) && \text{for } U \subseteq S, \\ &\lambda \leq x(S) \leq \kappa, \end{aligned}$$

is box-TDI.

Proof. Redefine $f_1(S) := \min\{f_1(S), \kappa\}$, and define $g_1 : \{S\} \to \mathbb{R}$ by $g_1(S) := \lambda$, and $g_2 : \emptyset \to \mathbb{R}$. Then (f_1, g_1) and (f_2, g_2) are paramodular pairs, and the box-total dual integrality of (49.94) is equivalent to the box-total dual integrality of (49.93). ∎

From Corollary 49.12c one can derive that the intersection of two integer generalized polymatroids is integer again. To prove this, we show that for any integer generalized polymatroid P there exists a paramodular pair (f, g) determining P, with f and g integer.

Let P be a generalized polymatroid, determined by the paramodular pair (f, g) of functions $f : \mathcal{C} \to \mathbb{R}$ and $g : \mathcal{D} \to \mathbb{R}$, where $\mathcal{C}$ and $\mathcal{D}$ are intersecting families. For any $U \subseteq S$, define

(49.95) $$\tilde{f}(U) := \max\{x(U) \mid x \in P\} \text{ and } \tilde{g}(U) := \min\{x(U) \mid x \in P\}.$$

So $\tilde{f}$ and $\tilde{g}$ are integer if P is integer.

Let

(49.96) $$\widetilde{\mathcal{C}} := \{U \in \mathcal{C} \mid \tilde{f}(U) < \infty\} \text{ and } \widetilde{\mathcal{D}} := \{U \in \mathcal{D} \mid \tilde{g}(U) > -\infty\}.$$

We restrict $\tilde{f}$ and $\tilde{g}$ to $\widetilde{\mathcal{C}}$ and $\widetilde{\mathcal{D}}$ respectively. We show that $(\tilde{f}, \tilde{g})$ is a paramodular pair determining P.

It is convenient to note that if $w \in \mathbb{R}^S$ with $w = w_1 + w_2$, then

(49.97) $$\max\{w^{\mathsf{T}}x \mid x \in P\} \leq \max\{w_1^{\mathsf{T}}x \mid x \in P\} + \max\{w_2^{\mathsf{T}}x \mid x \in P\}.$$

Theorem 49.13. *For any generalized polymatroid P, $\tilde{\mathcal{C}}$ and $\tilde{\mathcal{D}}$ are intersecting families, and the pair $(\tilde{f}, \tilde{g})$ is paramodular and determines P.*

Proof. We first show the following. Let $w \in \mathbb{Z}^S$ and let $\lambda > 0$ be such that $w_s \leq \lambda$ for each $s \in S$. Let $U := \{s \in S \mid w(s) = \lambda\}$ and $w' := w - \chi^U$. Then

(49.98) $$\max\{w^{\mathsf{T}}x \mid x \in P\} = \max\{w'^{\mathsf{T}}x \mid x \in P\} + \tilde{f}(U).$$

Here $\leq$ follows from (49.97), by definition of $\tilde{f}$. Equality is proved by induction on $|U|$, the case $U = \emptyset$ being trivial; so let $U \neq \emptyset$.

Let y, z be an optimum solution to the dual of $\max\{w^{\mathsf{T}}x \mid x \in P\}$:

(49.99) $$\min\{\sum_{T \in \mathcal{C}} y_T f(T) - \sum_{T \in \mathcal{D}} z_T g(T) \mid y \in \mathbb{R}^{\mathcal{C}}_+, z \in \mathbb{R}^{\mathcal{D}}_+, \sum_{T \in \mathcal{C}} y_T \chi^T - \sum_{T \in \mathcal{D}} z_T \chi^T = w\}.$$

Define $\mathcal{F} := \{T \in \mathcal{C} \mid y_T > 0\}$ and $\mathcal{G} := \{T \in \mathcal{D} \mid z_T > 0\}$. Similarly to Theorem 49.3, we can assume that $\mathcal{F} \cup \mathcal{G}$ is laminar.

Choose $u \in U$, and let W be an inclusionwise minimal set in $\mathcal{F}$ containing u. (Such a set exists, as $w(s) = \lambda > 0$.) Let $\mathcal{H}$ be the collection of inclusionwise maximal sets in $\mathcal{G}$ contained in $W - u$. As $\mathcal{G}$ is laminar, the sets in $\mathcal{H}$ are disjoint. Moreover, each $t \in W \setminus U$ is contained in some set in $\mathcal{H}$: since $w(t) < w(u)$ and since every set in $\mathcal{F}$ containing u also contains t (as $t \in W$), there exists an $X \in \mathcal{G}$ with $t \in X$ and $u \notin X$; as $\mathcal{F} \cup \mathcal{G}$ is laminar, we know that $X \subseteq W - u$.

Now let $Y := W \setminus \bigcup \mathcal{H}$. So Y is a nonempty subset of U. Define $w'' := w - \chi^Y$, let y' be obtained from y by decreasing $y(W)$ by 1, and let z' be obtained from z by decreasing $y(H)$ by 1 for each $H \in \mathcal{H}$. So (since $\chi^Y = \chi^W - \sum_{H \in \mathcal{H}} \chi^H$)

(49.100) $$\sum_{T\in\mathcal{C}} y'(T)\chi^T - \sum_{T\in\mathcal{D}} z'(T)\chi^T = w''$$

and

(49.101) $$\tilde{f}(Y) \leq f(W) - \sum_{H\in\mathcal{H}} g(H).$$

Moreover, setting $U' := U \setminus Y$, we have (by (49.97))

(49.102) $$\tilde{f}(U) \leq \tilde{f}(Y) + \tilde{f}(U'),$$

and by our induction hypothesis, as $|U'| < |U|$,

(49.103) $$\max\{w''^{\mathsf{T}}x \mid x \in P\} = \max\{w'^{\mathsf{T}}x \mid x \in P\} + \tilde{f}(U').$$

Hence

(49.104) $$\begin{aligned}\max\{w^{\mathsf{T}}x \mid x \in P\} &= \sum_{T\in\mathcal{C}} y(T)f(T) - \sum_{T\in\mathcal{D}} z(T)g(T)\\ &= \sum_{T\in\mathcal{C}} y'(T)f(T) - \sum_{T\in\mathcal{D}} z'(T)g(T) + f(W) - \sum_{H\in\mathcal{H}} g(H)\\ &\geq \max\{w''^{\mathsf{T}}x \mid x \in P\} + \tilde{f}(Y) = \max\{w'^{\mathsf{T}}x \mid x \in P\} + \tilde{f}(U') + \tilde{f}(Y)\\ &\geq \max\{w'^{\mathsf{T}}x \mid x \in P\} + \tilde{f}(U),\end{aligned}$$

thus proving (49.98).

We next derive that $\tilde{f}$ is submodular on intersecting pairs. Choose $X, Y \in \widetilde{\mathcal{C}}$ with $X \cap Y \neq \emptyset$. Define $w := \chi^X + \chi^Y$. Then by (49.98) and (49.97),

(49.105) $$\tilde{f}(X \cap Y) + \tilde{f}(X \cup Y) = \max\{w^{\mathsf{T}}x \mid x \in P\} \leq \tilde{f}(X) + \tilde{f}(Y).$$

So $\tilde{\mathcal{C}}$ is an intersecting family and $\tilde{f}$ is submodular on intersecting pairs. By symmetry, it follows that $\tilde{\mathcal{D}}$ is an intersecting family and $\tilde{g}$ is supermodular on intersecting pairs.

Finally, to see that (f, g) is paramodular, let $X \in \mathcal{C}$ and $Y \in \mathcal{D}$. Define $w := \chi^X - \chi^Y$. Again, by (49.98) and (49.97),

(49.106) $$\tilde{f}(X \setminus Y) - \tilde{g}(Y \setminus X) = \max\{w^{\mathsf{T}}x \mid x \in P\} \leq \tilde{f}(X) - \tilde{g}(Y).$$

So $\tilde{\mathcal{C}}$ and $\tilde{\mathcal{D}}$ are intersecting families, and the pair $(\tilde{f}, \tilde{g})$ is paramodular. It determines P, since P is determined by upper and lower bounds on $x(U)$ for subsets U of S. ∎

Corollary 49.12a implies:

Corollary 49.13a. *A generalized polymatroid P is integer if and only if there is a paramodular pair (f, g) defining P with f and g integer.*

Proof. Sufficiency follows from Corollary 49.12a. Necessity follows from Theorem 49.13, as P is determined by $(\tilde{f}, \tilde{g})$, where $\tilde{f}$ and $\tilde{g}$ are integer if P is integer. ∎

A second consequence is:

Corollary 49.13b. *The intersection of two integer generalized polymatroids is integer.*

Proof. Directly by combining Corollaries 49.12c and 49.13a. ∎

We should note that the collections $\tilde{\mathcal{C}}$ and $\tilde{\mathcal{D}}$ found in the proof of Theorem 49.13 are lattice families and that $\tilde{f}$ and $\tilde{g}$ are sub- and supermodular respectively. Moreover, $\tilde{f}(T \setminus U) - \tilde{g}(U \setminus T) \leq \tilde{f}(T) - \tilde{g}(U)$ for each pair $T \in \tilde{\mathcal{C}}$, $U \in \tilde{\mathcal{D}}$.

This implies that projections of generalized polymatroids are again generalized polymatroids (Frank [1984b]):

Corollary 49.13c. *Let $P \subseteq \mathbb{R}^S$ be a generalized polymatroid and let $t \in S$. Define $S' := S - t$. Then the projection*

(49.107) $$P' := \{x \in \mathbb{R}^{S'} \mid \exists \eta : (x, \eta) \in P\}$$

is again a generalized polymatroid.

Proof. We can assume that P is nonempty, that $\mathcal{C}$ and $\mathcal{D}$ are lattice families, and that P is determined by a paramodular pair $(f, g) = (\tilde{f}, \tilde{g})$ as above. Let $\mathcal{C}'$ and $\mathcal{D}'$ be the collections of sets in $\mathcal{C}$ and $\mathcal{D}$ respectively not containing t. Let $f' := f|\mathcal{C}'$ and $g' := g|\mathcal{D}'$.

Trivially, (f', g') is a paramodular pair. We claim that P' is equal to the generalized polymatroid Q determined by (f', g'). Trivially, $P' \subseteq Q$. To see the reverse inclusion, let $x \in Q$. Let η' be the largest real such that $x(T - t) + \eta' \leq f(T)$ for each $T \in \mathcal{C} \setminus \mathcal{C}'$. Let η'' be the smallest real such that $x(U - t) + \eta'' \geq g(U)$ for each $U \in \mathcal{D} \setminus \mathcal{D}'$.

If $x \notin P'$, then $\eta' < \eta''$, and hence there exist $T \in \mathcal{C}$ and $U \in \mathcal{D}$ with $t \in T \cap U$ and $f(T) - x(T - t) < g(U) - x(U - t)$. Hence

(49.108) $$\begin{aligned} x(T \setminus U) - x(U \setminus T) &= x(T - t) - x(U - t) > f(T) - g(U) \\ &\geq f(T \setminus U) - g(U \setminus T). \end{aligned}$$

This contradicts the fact that $x(T \setminus U) \leq f(T \setminus U)$ and $x(U \setminus T) \geq g(U \setminus T)$, as $x \in Q$. ∎

For results on the dimension of generalized polymatroids, see Frank and Tardos [1988], which paper surveys generalized polymatroids and submodular flows. More results on generalized polymatroids are reported by Fujishige [1984b], Nakamura [1988b], Naitoh and Fujishige [1992], and Tamir [1995].

49.11c. Supermodular colourings

A colouring-type of result on supermodular functions was shown by Schrijver [1985]. We give the proof based on generalized polymatroids found by Tardos [1985b].

Theorem 49.14. *Let $\mathcal{C}_1$ and $\mathcal{C}_2$ be intersecting families of subsets of a set S, let $g_1 : \mathcal{C}_1 \to \mathbb{Z}$ and $g_2 : \mathcal{C}_2 \to \mathbb{Z}$ be intersecting supermodular, and let $k \in \mathbb{Z}_+$ with $k \geq 1$. Then S can be partitioned into classes $L_1, \ldots, L_k$ such that*

(49.109) $$g_i(U) \leq |\{j \in \{1, \ldots, k\} \mid L_j \cap U \neq \emptyset\}|$$

for each $i = 1, 2$ and each $U \in \mathcal{C}_i$ if and only if

(49.110) $$g_i(U) \leq \min\{k, |U|\}$$

for each $i = 1, 2$ *and each* $U \in \mathcal{C}_i$.

Proof. Necessity is easy. Sufficiency is shown by induction on k, the case $k = 0$ being trivial. By induction, it suffices to find a subset L of S such that

(49.111) $$|U \setminus L| \geq g_i(U) - 1 \text{ and, if } g_i(U) = k, \text{ then } U \cap L \neq \emptyset.$$

Indeed, in that case we can apply induction to the functions $g_i' : \mathcal{C}_i' \to \mathbb{Z}$ on $\mathcal{C}_i' := \{U \setminus L \mid U \in \mathcal{C}\}$, defined by

(49.112) $$g_i'(U \setminus L) := \begin{cases} g_i(U) - 1 & \text{if } U \cap L \neq \emptyset, \\ g_i(U) & \text{if } U \cap L = \emptyset, \end{cases}$$

for $U \in \mathcal{C}_i$.

For $i = 1, 2$, consider the system:

(49.113) $$\begin{array}{rll} \text{(i)} & 0 \leq x_s \leq 1 & \text{for } s \in S, \\ \text{(ii)} & x(U) \leq |U| - g_i(U) + 1 & \text{for } U \in \mathcal{C}_i, \\ \text{(iii)} & x(U) \geq 1 & \text{for } U \in \mathcal{C}_i \text{ with } g_i(U) = k. \end{array}$$

This system determines an integer generalized polymatroid. This can be seen as follows. Let $\mathcal{D}_i$ be the collection of inclusionwise minimal sets in $\{U \in \mathcal{C}_i \mid g_i(U) = k\}$. So $\mathcal{D}_i$ consists of disjoint sets (as $\mathcal{C}_i$ is intersecting and as $g_i(U) \leq k$ for each $U \in \mathcal{C}_i$). Let

(49.114) $$\mathcal{C}_i' := \{U \in \mathcal{C}_i \mid \forall T \in \mathcal{D}_i : U \subseteq T \text{ or } T \cap U = \emptyset\}.$$

Then (49.113) has the same solution set as:

(49.115) $$\begin{array}{rll} \text{(i)} & 0 \leq x_s \leq 1 & \text{for } s \in S, \\ \text{(ii)} & x(U) \leq |U| - g_i(U) + 1 & \text{for } U \in \mathcal{C}_i', \\ \text{(iii)} & x(U) \geq 1 & \text{for } U \in \mathcal{D}_i. \end{array}$$

Indeed, (49.115)(iii) implies (49.113)(iii) (as $x \geq \mathbf{0}$). Moreover, for any $U \in \mathcal{C}_i$ with $T \cap U \neq \emptyset$ for some $T \in \mathcal{D}_i$, one has

(49.116) $$g_i(T \cap U) \geq g_i(T) + g_i(U) - g_i(T \cup U) \geq g_i(U)$$

(as $g_i(T \cup U) \leq k = g_i(T)$). So with (49.115)(iii) we have:

(49.117) $$x(U) \leq x(T \cap U) + |U \setminus T| \leq |T \cap U| - g_i(T \cap U) + 1 + |U \setminus T| \\ \leq |U| - g_i(U) + 1$$

(as $x_s \leq 1$ for all $x \in U \setminus T$). Hence (49.113) and (49.115) have the same solution set.

Now (49.115) is a system defining a generalized polymatroid, as one easily checks (condition (49.87)(iii) follows, since if $T \in \mathcal{C}_i'$ and $U \in \mathcal{D}_i$ with $T \setminus U \neq \emptyset$ and $U \setminus T \neq \emptyset$, then, by definition of $\mathcal{C}_i'$, T and U are disjoint, and then the inequality in (49.87)(iii) is trivial). It is integer, as the right-hand sides in (49.115) are integer.

Also, the intersection of these generalized polymatroids for $i = 1$ and $i = 2$ is nonempty, since the vector $x := k^{-1} \cdot \mathbf{1}$ belongs to it. (49.115)(i) and (iii) hold trivially. To see (ii), we have

(49.118) $$x(U) = \tfrac{1}{k}|U| = |U| - \tfrac{k-1}{k}|U| \leq |U| - \tfrac{k-1}{k} g_i(U) = |U| - g_i(U) + \tfrac{1}{k} g_i(U) \\ \leq |U| - g_i(U) + 1.$$

Therefore, the intersection contains an integer vector x, which is, by (49.115)(i), the incidence vector of some subset L of S satisfying (49.111), as required. ∎

This theorem generalizes edge-colouring theorems for bipartite graphs $G = (V, E)$. Let V_1 and V_2 be the colour classes of G. Let $\mathcal{C}_i := \{\delta(v) \mid v \in V_i\}$ for $i = 1, 2$. If we define $g_i(\delta(v)) := |\delta(v)|$ for $v \in V_i$ $(i = 1, 2)$, Theorem 49.14 reduces to Kőnig's edge-colouring theorem (Theorem 20.1). If $g_i(\delta(v))$ is set to the minimum degree of G, we obtain Theorem 20.5, and if it is set to the minimum of k and $|\delta(v)|$, we obtain Theorem 20.6.

Theorem 49.14 can also be used in proving the 'disjoint bibranchings theorem' (Theorem 54.11 — see Section 54.7a). Szigeti [1999] gave a generalization of Theorem 49.14.

49.11d. Further notes

Let S and T be disjoint sets. A function $f : \mathcal{P}(S) \times \mathcal{P}(T) \to \mathbb{R}$ is called *bisubmodular* if

$$(49.119) \qquad f(X_1 \cap X_2, Y_1 \cup Y_2) + f(X_1 \cup X_2, Y_1 \cap Y_2) \leq f(X_1, Y_1) + f(X_2, Y_2)$$

for all $X_1, X_2 \subseteq S$ and $Y_1, Y_2 \subseteq T$.

Bisubmodular functions were studied by Kung [1978b] and Schrijver [1978, 1979c]. Most of the results can be obtained from those for submodular functions, by considering the submodular set function f' on $S \cup T$ defined by $f'(X \cup Y) := f((S \setminus X) \cup Y)$ for $X \subseteq S$ and $Y \subseteq T$. Similarly for *bisupermodular* functions, where the inequality sign in (49.119) is reversed.

For an interesting related result of Frank and Jordán [1995b] yielding Győri's theorem, see Section 60.3d.

Fujishige [1984c] gave a framework that includes Theorem 46.2 on the total dual integrality of the intersection of a polymatroid and a contrapolymatroid system, Corollary 46.2b on the existence of a modular function between a sub- and a supermodular function, and Theorem 49.13 on the total dual integrality of the generalized polymatroid constraints (but not the total dual integrality of the intersection of two polymatroids). Fujishige [1984b] described generalized polymatroids as projections of base polyhedra of submodular functions.

Chandrasekaran and Kabadi [1988] introduced the concept of a *generalized submodular function* as a function $f : \mathcal{R} \to \mathbb{R}$, where $\mathcal{R} := \{(T, U) \mid T, U \subseteq S, T \cap U = \emptyset\}$ for some set S, satisfying

$$(49.120) \qquad \begin{aligned} & f(A, B) + f(C, D) \\ & \geq f(A \cap C, B \cap D) + f((A \setminus D) \cup (C \setminus B), (B \setminus C) \cup (D \setminus A)) \end{aligned}$$

for all $(A, B), (C, D) \in \mathcal{R}$. They showed that the system

$$(49.121) \qquad x(T) - x(U) \leq f(T, U) \text{ for } (T, U) \in \mathcal{R}$$

is box-TDI, and that for any $w \in \mathbb{R}^S$, an x maximizing $w^{\mathsf{T}} x$ over (49.121) can be found by a variant of the greedy method. Unions of two such systems need not define an integer polyhedron if the functions are integer, as is shown by an example with $|S| = 2$. A similar framework was considered by Nakamura [1990]. More results can be found in Dress and Havel [1986], Bouchet [1987a,1995], Bouchet, Dress, and

Havel [1992], Ando and Fujishige [1996], Fujishige [1997], and Fujishige and Iwata [2001].

It is direct to represent a lattice family on a set S of size n in $O(n^2)$ space (just by giving all pairs (u, v) for which each set in the family containing u also contains v). Gabow [1993b,1995c] gave an $O(n^2)$ representation for intersecting and crossing families. Related results were found by Fleiner and Jordán [1999].

Tardos [1985b] also studied *generalized matroids*, which form the special case of generalized polymatroids with $0, 1$ vertices. An instance of it we saw in the proof of Theorem 49.14.

More results on submodularity are given by Fujishige [1980b,1984f,1984g,1988], Nakamura [1988b,1988c,1993], Kabadi and Chandrasekaran [1990], Iwata [1995], Iwata, Murota, and Shigeno [1997], and Murota [1998]. Generalizations were studied by Qi [1988b] and Kashiwabara, Nakamura, and Takabatake [1999].