Prof. Jeff Bilmes

University of Washington, Seattle
Department of Electrical Engineering
Spring Quarter, 2011
http://ssli.ee.washington.edu/~bilmes/ee595a_spring_2011/

Lecture 17 - May 27th, 2011
Announcements

- Last lecture, and final presentations, will take place Thursday, June 9th, from 3-7:30pm. The lecture will be from 3:00-5:00pm, and the final presentations will be from 5:00-7:30pm. Please bring dinner.
We need to find one makeup lecture this term.

- L1 (3/30):
- L2 (4/1):
- L3 (4/6):
- L4 (4/8):
- L5 (4/13):
- L6 (4/15):
- L7 (4/20):
- L8 (4/27):
- L9 (4/29):
- L10 (5/4):
- L13 (5/13): More polymatroids, start lattices
- L14 (5/18): lattices/submodular
- L15 (5/20): lattices, \( \rightarrow \) SFM.
- L16 (5/25): \( \rightarrow \) SFM
- L17 (5/27):
- L18 (6/1):
- L19 (6/3):
- L20: (6/9): 3-7:30pm (EEB-303)?
Consider $x \in P_f$, and consider the following set

$$
\text{DEP}(x) = \{ \text{dep}(x, e) : e \in \text{sat}(x) \} \quad (1)
$$
Consider $x \in P_f$, and consider the following set
\[ \text{DEP}(x) = \{ \text{dep}(x, e) : e \in \text{sat}(x) \} \] (1)

So DEP($x$) is a set of sets, each element of DEP($x$) is the dep($x$, $e$) valuation for some $e \in \text{sat}(x)$.
Consider \( x \in P_f \), and consider the following set
\[
\text{DEP}(x) = \{ \text{dep}(x, e) : e \in \text{sat}(x) \} \tag{1}
\]
So \( \text{DEP}(x) \) is a set of sets, each element of \( \text{DEP}(x) \) is the \( \text{dep}(x, e) \) valuation for some \( e \in \text{sat}(x) \).

Moreover, define a partial order on \( \text{DEP}(x) \) as follows: if \( A, B \in \text{DEP}(x) \), then \( A \preceq B \) iff \( A \subseteq B \).
Consider $x \in P_f$, and consider the following set

$$\text{DEP}(x) = \{\text{dep}(x, e) : e \in \text{sat}(x)\}$$ (1)

So $\text{DEP}(x)$ is a set of sets, each element of $\text{DEP}(x)$ is the $\text{dep}(x, e)$ valuation for some $e \in \text{sat}(x)$.

Moreover, define a partial order on $\text{DEP}(x)$ as follows: if $A, B \in \text{DEP}(x)$, then $A \preceq B$ iff $A \subseteq B$.

We’re going to use this partial order to define a partial order on all elements of $\text{sat}(x)$. 
Consider $x \in P_f$, and consider the following set
\[ \text{DEP}(x) = \{ \text{dep}(x, e) : e \in \text{sat}(x) \} \]
So $\text{DEP}(x)$ is a set of sets, each element of $\text{DEP}(x)$ is the $\text{dep}(x, e)$ valuation for some $e \in \text{sat}(x)$.
Moreover, define a partial order on $\text{DEP}(x)$ as follows: if $A, B \in \text{DEP}(x)$, then $A \preceq B$ iff $A \subseteq B$.
We’re going to use this partial order to define a partial order on all elements of $\text{sat}(x)$.
Now recall $\mathcal{D}(x) = \{ A : x(A) = f(A) \}$ forms a distributive lattice. What is the natural partial order?
Now in any distributive lattice $L$, consider its join-irreducibles $\mathcal{J}$ (i.e., any element $A \in \mathcal{J}$ can’t be represented as a join of any other two elements in $L$).

We saw that if the lattice has length $n$, then $\mathcal{J}$ will have exactly $n$ elements (in the Boolean case, these are atoms/ground elements), and each element in $\mathcal{J}$ is partially ordered by the lattice partial order.

Moreover, we saw any element can be “generated” by joining the join-irreducible elements.
Now any element in \( \text{DEP}(x) \) (for \( x \) extreme) can’t be represented by the join of two other elements in \( \text{DEP}(x) \), since the minimal tight sets containing \( e \) would not be generated by merging two minimal tight sets containing, say, \( a \), and \( b \), where all of \( a, b, e \) are unequal.

Thus, considering \( \mathcal{D}(x) \) as a distributed lattice, then \( \text{DEP}(x) \) are the join-irreducibles.

And the order \( \preceq \) defined earlier is the natural order w.r.t. this lattice and its join-irreducibles.
Let \( x \in P_f \) again be an extreme point, and let it be generated by an ordering of \( B = (e_1, e_2, \ldots, e_k) \subseteq E \) with \( B_i = (b_i, b_2, \ldots, b_i) \) a partial order w.r.t. ordered items \( B \) (\( B \) and \( B_i, \forall i \) are ordered sets).
dep and partial order

Let \( x \in P_f \) again be an extreme point, and let it be generated by an ordering of \( B = (e_1, e_2, \ldots, e_k) \subseteq E \) with \( B_i = (b_i, b_2, \ldots, b_i) \) a partial order w.r.t. ordered items \( B \) (\( B \) and \( B_i, \forall i \) are ordered sets).

Recall, the equation for \( x \) is of the form \( x(e) = 0 \) for some \( e \) and \( x(A) = f(A) \) for some \( A \) (see earlier). Specifically, we have that \( x(E \setminus B) = 0 \) and, for \( i = 1 \ldots k \), \( x(B_i) = f(B_i) \).
Let $x \in P_f$ again be an extreme point, and let it be generated by an ordering of $B = (e_1, e_2, \ldots, e_k) \subseteq E$ with $B_i = (b_i, b_2, \ldots, b_i)$ a partial order w.r.t. ordered items $B$ ($B$ and $B_i, \forall i$ are ordered sets).

Recall, the equation for $x$ is of the form $x(e) = 0$ for some $e$ and $x(A) = f(A)$ for some $A$ (see earlier). Specifically, we have that $x(E \setminus B) = 0$ and, for $i = 1 \ldots k$, $x(B_i) = f(B_i)$.

Thus, each of $B_i$ is a tight set.
dep and partial order

Let \( x \in P_f \) again be an extreme point, and let it be generated by an ordering of \( B = (e_1, e_2, \ldots, e_k) \subseteq E \) with \( B_i = (b_i, b_2, \ldots, b_i) \) a partial order w.r.t. ordered items \( B \) (\( B \) and \( B_i, \forall i \) are ordered sets).

Recall, the equation for \( x \) is of the form \( x(e) = 0 \) for some \( e \) and \( x(A) = f(A) \) for some \( A \) (see earlier). Specifically, we have that \( x(E \setminus B) = 0 \) and, for \( i = 1 \ldots k \), \( x(B_i) = f(B_i) \).

Thus, each of \( B_i \) is a tight set.

We also have that \( \text{supp}(x) \subseteq B \) due to monotonicity.
Let \( x \in P_f \) again be an extreme point, and let it be generated by an ordering of \( B = (e_1, e_2, \ldots, e_k) \subseteq E \) with \( B_i = (b_i, b_2, \ldots, b_i) \) a partial order w.r.t. ordered items \( B \) (\( B \) and \( B_i, \forall i \) are ordered sets).

Recall, the equation for \( x \) is of the form \( x(e) = 0 \) for some \( e \) and \( x(A) = f(A) \) for some \( A \) (see earlier). Specifically, we have that \( x(E \setminus B) = 0 \) and, for \( i = 1 \ldots k \), \( x(B_i) = f(B_i) \).

Thus, each of \( B_i \) is a tight set.

We also have that \( \text{supp}(x) \subseteq B \) due to monotonicity.

Thus, for any \( d, e \in \text{supp}(x) \subseteq B \), there is a tight set containing one but not the other. Specifically, let \( d = e_i \) and \( e = e_j \) with \( j > i \). Then non-zero \( B_i \) (i.e., \( B_i \cap \text{supp}(x) \)) contains \( d \) but not \( e \) (note, vice versa is not true).
Let \( x \in P_f \) again be an extreme point, and let it be generated by an ordering of \( B = (e_1, e_2, \ldots, e_k) \subseteq E \) with \( B_i = (b_i, b_2, \ldots, b_i) \) a partial order w.r.t. ordered items \( B \) (\( B \) and \( B_i, \forall i \) are ordered sets).

Recall, the equation for \( x \) is of the form \( x(e) = 0 \) for some \( e \) and \( x(A) = f(A) \) for some \( A \) (see earlier). Specifically, we have that \( x(E \setminus B) = 0 \) and, for \( i = 1 \ldots k \), \( x(B_i) = f(B_i) \).

Thus, each of \( B_i \) is a tight set.

We also have that \( \text{supp}(x) \subseteq B \) due to monotonicity.

Thus, for any \( d, e \in \text{supp}(x) \subseteq B \), there is a tight set containing one but not the other. Specifically, let \( d = e_i \) and \( e = e_j \) with \( j > i \). Then non-zero \( B_i \) (i.e., \( B_i \cap \text{supp}(x) \)) contains \( d \) but not \( e \) (note, vice versa is not true).

Thus, for any \( d, e \in \text{supp}(x) \subseteq B \), we have \( \text{dep}(x, d) \neq \text{dep}(x, e) \).
dep and partial order

- Let \( x \in P_f \) again be an extreme point, and let it be generated by an ordering of \( B = (e_1, e_2, \ldots, e_k) \subseteq E \) with \( B_i = (b_i, b_2, \ldots, b_i) \) a partial order w.r.t. ordered items \( B \) (\( B \) and \( B_i, \forall i \) are ordered sets).

- Recall, the equation for \( x \) is of the form \( x(e) = 0 \) for some \( e \) and \( x(A) = f(A) \) for some \( A \) (see earlier). Specifically, we have that \( x(E \setminus B) = 0 \) and, for \( i = 1 \ldots k \), \( x(B_i) = f(B_i) \).

- Thus, each of \( B_i \) is a tight set.

- We also have that \( \text{supp}(x) \subseteq B \) due to monotonicity.

- Thus, for any \( d, e \in \text{supp}(x) \subseteq B \), there is a tight set containing one but not the other. Specifically, let \( d = e_i \) and \( e = e_j \) with \( j > i \). Then non-zero \( B_i \) (i.e., \( B_i \cap \text{supp}(x) \)) contains \( d \) but not \( e \) (note, vice versa is not true).

- Thus, for any \( d, e \in \text{supp}(x) \subseteq B \), we have \( \text{dep}(x, d) \neq \text{dep}(x, e) \).

- Moreover, for any \( e \in B \), we can have that \( \text{dep}(x, e) = B_i \) where \( e = e_i \). This point is further clarified in the next slide.
dep and partial order (slight digression)

- I.e., $x$ is extreme generated by $B$, then $B_i$ is a tight set containing $e_i$. 
I.e., $x$ is extreme generated by $B$, then $B_i$ is a tight set containing $e_i$. For any $j < i$, $B_j$ does not contain $e_i$. 
dep and partial order (slight digression)

- I.e., \( x \) is extreme generated by \( B \), then \( B_i \) is a tight set containing \( e_i \).
- For any \( j < i \), \( B_j \) does not contain \( e_i \).
- Thus, \( \text{dep}(x, e_i) \) (minimal tight \( e_i \)-containing set) might equal \( B_i \).
dep and partial order (slight digression)

- I.e., $x$ is extreme generated by $B$, then $B_i$ is a tight set containing $e_i$.
- For any $j < i$, $B_j$ does not contain $e_i$.
- Thus, $\text{dep}(x, e_i)$ (minimal tight $e_i$-containing set) might equal $B_i$.
- On the other hand, consider the extreme vector $x^{(i)} \in \mathbb{R}^E$ with

$$x^{(i)}(e) = \begin{cases} x(e) & \text{if } e \in B_i \\ 0 & \text{else} \end{cases}$$  \(2\)

so $x^{(i)}$ is just the extreme vector generated by the ordered set $B_i$. 
dep and partial order (slight digression)

- I.e., $x$ is extreme generated by $B$, then $B_i$ is a tight set containing $e_i$.
- For any $j < i$, $B_j$ does not contain $e_i$.
- Thus, $\text{dep}(x, e_i)$ (minimal tight $e_i$-containing set) might equal $B_i$.
- On the other hand, consider the extreme vector $x^{(i)} \in \mathbb{R}^E$ with

$$x^{(i)}(e) = \begin{cases} 
    x(e) & \text{if } e \in B_i \\
    0 & \text{else}
\end{cases}$$

so $x^{(i)}$ is just the extreme vector generated by the ordered set $B_i$.
- Therefore, $B_j$ for $j \leq i$ are tight w.r.t. $x^{(i)}$. 
dep and partial order (slight digression)

- I.e., $x$ is extreme generated by $B$, then $B_i$ is a tight set containing $e_i$.
- For any $j < i$, $B_j$ does not contain $e_i$.
- Thus, $\text{dep}(x, e_i)$ (minimal tight $e_i$-containing set) might equal $B_i$.
- On the other hand, consider the extreme vector $x^{(i)} \in \mathbb{R}^E$ with
  \[
  x^{(i)}(e) = \begin{cases} 
  x(e) & \text{if } e \in B_i \\
  0 & \text{else}
  \end{cases}
  \] (2)
  so $x^{(i)}$ is just the extreme vector generated by the ordered set $B_i$.
- Therefore, $B_j$ for $j \leq i$ are tight w.r.t. $x^{(i)}$.
- Could be other ordered sets (say $B^{(i)}$, which is $B_i$ permuted) that also generates $x^{(i)}$. Let $B^{(i)}_j, j \leq i$ be the first $j$ elements in $B^{(i)}$. 
dep and partial order (slight digression)

- I.e., $x$ is extreme generated by $B$, then $B_i$ is a tight set containing $e_i$.
- For any $j < i$, $B_j$ does not contain $e_i$.
- Thus, $\text{dep}(x, e_i)$ (minimal tight $e_i$-containing set) might equal $B_i$.
- On the other hand, consider the extreme vector $x^{(i)} \in \mathbb{R}^E$ with
  
  \[
  x^{(i)}(e) = \begin{cases} 
  x(e) & \text{if } e \in B_i \\
  0 & \text{else}
  \end{cases}
  \]

  so $x^{(i)}$ is just the extreme vector generated by the ordered set $B_i$.
- Therefore, $B_j$ for $j \leq i$ are tight w.r.t. $x^{(i)}$.
- Could be other ordered sets (say $B^{(i)}$, which is $B_i$ permuted) that also generates $x^{(i)}$. Let $B^{(i)}_j$, $j \leq i$ be the first $j$ elements in $B^{(i)}$.
- In $B^{(i)}$, $e_i$ might come at a position $j < i$, so $B^{(i)}_j$ is tight and containing $e_i$, and $\text{dep}(x, e_i)$ might equal $B^{(i)}_j$, with $B^{(i)}_j \subset B_i$. 
dep and partial order (slight digression)

- I.e., \( x \) is extreme generated by \( B \), then \( B_i \) is a tight set containing \( e_i \).
- For any \( j < i \), \( B_j \) does not contain \( e_i \).
- Thus, \( \text{dep}(x, e_i) \) (minimal tight \( e_i \)-containing set) might equal \( B_i \).
- On the other hand, consider the extreme vector \( x^{(i)} \in \mathbb{R}^E \) with
  \[
  x^{(i)}(e) = \begin{cases} 
  x(e) & \text{if } e \in B_i \\
  0 & \text{else}
  \end{cases}
  \]  
  (2)
  so \( x^{(i)} \) is just the extreme vector generated by the ordered set \( B_i \).
- Therefore, \( B_j \) for \( j \leq i \) are tight w.r.t. \( x^{(i)} \).
- Could be other ordered sets (say \( B^{(i)} \), which is \( B_i \) permuted) that also generates \( x^{(i)} \). Let \( B^{(i)}_j \), \( j \leq i \) be the first \( j \) elements in \( B^{(i)} \).
- In \( B^{(i)} \), \( e_i \) might come at a position \( j < i \), so \( B^{(i)}_j \) is tight and containing \( e_i \), and \( \text{dep}(x, e_i) \) might equal \( B^{(i)}_j \), with \( B^{(i)}_j \subset B_i \).
- On the other hand, \( B_i \not\subset \text{dep}(x, e_i) \) due to \( \text{dep}(x, e_i) \)’s minimality.
Dep and partial order (slight digression)

- I.e., \( x \) is extreme generated by \( B \), then \( B_i \) is a tight set containing \( e_i \).
- For any \( j < i \), \( B_j \) does not contain \( e_i \).
- Thus, \( \text{dep}(x, e_i) \) (minimal tight \( e_i \)-containing set) might equal \( B_i \).
- On the other hand, consider the extreme vector \( x^{(i)} \in \mathbb{R}^E \) with
  \[
  x^{(i)}(e) = \begin{cases} 
  x(e) & \text{if } e \in B_i \\
  0 & \text{else}
  \end{cases}
  \]  
  \[ (2) \]
  so \( x^{(i)} \) is just the extreme vector generated by the ordered set \( B_i \).
- Therefore, \( B_j \) for \( j \leq i \) are tight w.r.t. \( x^{(i)} \).
- Could be other ordered sets (say \( B^{(i)} \), which is \( B_i \) permuted) that also generates \( x^{(i)} \). Let \( B_j^{(i)} \), \( j \leq i \) be the first \( j \) elements in \( B^{(i)} \).
- In \( B^{(i)} \), \( e_i \) might come at a position \( j < i \), so \( B_j^{(i)} \) is tight and containing \( e_i \), and \( \text{dep}(x, e_i) \) might equal \( B_j^{(i)} \), with \( B_j^{(i)} \subset B_i \).
- On the other hand, \( B_i \not\subset \text{dep}(x, e_i) \) due to \( \text{dep}(x, e_i) \)'s minimality.
- Therefore, we see that in general, \( \text{dep}(x, e_i) \subset B_i \).
dep and partial order (slight digression)

Now, while $\text{dep}(x, e_i) \subseteq B_i$, we can be a bit more explicit.
dep and partial order (slight digression)

- Now, while $\text{dep}(x, e_i) \subseteq B_i$, we can be a bit more explicit.
- Let $B(x)$ be set of permutations of $B$ that generate $x$. 
dep and partial order (slight digression)

- Now, while $\text{dep}(x, e_i) \subseteq B_i$, we can be a bit more explicit.
- Let $B(x)$ be set of permutations of $B$ that generate $x$.
- For $e \in B$ and $B' \in B(x)$, let $1 \leq e(B') \leq |B'|$ be $e$'s position in $B'$.
Now, while \( \text{dep}(x, e_i) \subseteq B_i \), we can be a bit more explicit.

Let \( B(x) \) be set of permutations of \( B \) that generate \( x \).

For \( e \in B \) and \( B' \in B(x) \), let \( 1 \leq e(B') \leq |B'| \) be \( e \)'s position in \( B' \).

Then \( \text{dep}(x, e_i) = B^e_i \) where

\[
B^e_i \in \arg\min_{B' \in B(x)} e_i(B') \quad \text{and also} \quad \left| \arg\min_{B' \in B(x)} e_i(B') \right| = 1 \quad (3)
\]

is ordered, and \( j \) is the position of \( e_i \) in \( B^e_i \). Follows from iff relationship between extremal points and greedy algorithm, and since \( \text{dep}(x, e_i) \) is the unique “0” element of a distributive lattice.
dep and partial order (slight digression)

- Now, while \( \text{dep}(x, e_i) \subseteq B_i \), we can be a bit more explicit.
- Let \( B(x) \) be set of permutations of \( B \) that generate \( x \).
- For \( e \in B \) and \( B' \in B(x) \), let \( 1 \leq e(B') \leq |B'| \) be \( e \)'s position in \( B' \).
- Then \( \text{dep}(x, e_i) = B^{e_i}_j \) where

\[
B^{e_i}_j \in \arg\min_{B' \in B(x)} e_i(B') \quad \text{and also} \quad \left| \arg\min_{B' \in B(x)} e_i(B') \right| = 1 \quad (3)
\]

is ordered, and \( j \) is the position of \( e_i \) in \( B^{e_i} \). Follows from iff relationship between extremal points and greedy algorithm, and since \( \text{dep}(x, e_i) \) is the unique “0” element of a distributive lattice.
- Note, for \( e \in \text{sat}(x) \), \( B^e_j \subseteq B \), and \( |B^e_j| = j \).
dep and partial order (slight digression)

- Now, while \( \text{dep}(x, e_i) \subseteq B_i \), we can be a bit more explicit.
- Let \( B(x) \) be set of permutations of \( B \) that generate \( x \).
- For \( e \in B \) and \( B' \in B(x) \), let \( 1 \leq e(B') \leq |B'| \) be \( e \)'s position in \( B' \).
- Then \( \text{dep}(x, e_i) = B_{e_i}^j \) where

\[
B_{e_i}^j \in \argmin_{B' \in B(x)} e_i(B') \quad \text{and also} \quad |\argmin_{B' \in B(x)} e_i(B')| = 1 \quad (3)
\]

is ordered, and \( j \) is the position of \( e_i \) in \( B_{e_i}^j \). Follows from iff relationship between extremal points and greedy algorithm, and since \( \text{dep}(x, e_i) \) is the unique “0” element of a distributive lattice.

- Note, for \( e \in \text{sat}(x) \), \( B^e_j \subseteq B \), and \( |B^e_j| = j \).
- Also, for \( d, e \in \text{sat}(x) \), \( \text{dep}(x, d) = B_d^i \subseteq \text{dep}(x, e) = B_e^j \) iff \( d \in \text{dep}(x, e) \).
dep and partial order (slight digression)

- Now, while $\text{dep}(x, e_i) \subseteq B_i$, we can be a bit more explicit.
- Let $B(x)$ be set of permutations of $B$ that generate $x$.
- For $e \in B$ and $B' \in B(x)$, let $1 \leq e(B') \leq |B'|$ be $e$’s position in $B'$.
- Then $\text{dep}(x, e_i) = B_j^{e_i}$ where
  
  $B_j^{e_i} \in \arg\min_{B' \in B(x)} e_i(B')$ and also $|\arg\min_{B' \in B(x)} e_i(B')| = 1$ (3)

  is ordered, and $j$ is the position of $e_i$ in $B_j^{e_i}$. Follows from iff relationship between extremal points and greedy algorithm, and since $\text{dep}(x, e_i)$ is the unique “0” element of a distributive lattice.

- Note, for $e \in \text{sat}(x)$, $B_j^e \subseteq B$, and $|B_j^e| = j$.

- Also, for $d, e \in \text{sat}(x)$, $\text{dep}(x, d) = B_i^d \subset \text{dep}(x, e) = B_j^e$ iff $d \in \text{dep}(x, e)$.
  - Clearly, $\text{dep}(x, d) \subset \text{dep}(x, e) \Rightarrow d \in \text{dep}(x, e)$. 
Now, while \( \text{dep}(x, e_i) \subseteq B_i \), we can be a bit more explicit.

Let \( B(x) \) be set of permutations of \( B \) that generate \( x \).

For \( e \in B \) and \( B' \in B(x) \), let \( 1 \leq e(B') \leq |B'| \) be \( e \)'s position in \( B' \).

Then \( \text{dep}(x, e_i) = B^e_i \) where

\[
B^e_i \in \arg\min_{B' \in B(x)} e_i(B') \quad \text{and also} \quad |\arg\min_{B' \in B(x)} e_i(B')| = 1 \quad (3)
\]

is ordered, and \( j \) is the position of \( e_i \) in \( B^e_i \). Follows from iff relationship between extremal points and greedy algorithm, and since \( \text{dep}(x, e_i) \) is the unique “0” element of a distributive lattice.

Note, for \( e \in \text{sat}(x) \), \( B^e_j \subseteq B \), and \( |B^e_j| = j \).

Also, for \( d, e \in \text{sat}(x) \), \( \text{dep}(x, d) = B^d_i \subseteq \text{dep}(x, e) = B^e_j \) iff \( d \in \text{dep}(x, e) \).

Clearly, \( \text{dep}(x, d) \subseteq \text{dep}(x, e) \Rightarrow d \in \text{dep}(x, e) \).

Also \( d \in \text{dep}(x, e) \) means \( \text{dep}(x, d) \subseteq B^e_k \) where \( k = d(B^e) \) is the position of \( d \) in \( B^e \) (since \( B^e_k \) is a tight set containing \( d \)), but it must be that \( k < j \) (since \( B^e_j \) is the smallest tight set containing \( e \) and the \( j \)'th position of \( B^e_j \) is \( e \)).
Also, for polymatroidal $f$, we saw earlier that for each $e \in \text{sat}(x) \setminus \text{supp}(x)$, the set $\text{supp}(x) + e$ is also tight. This follows since $x(\text{supp}(x) + e) = x(\text{supp}(x))$ but $e$ is dependent on $\text{supp}(x)$ so that $f(\text{supp}(x) + e) = f(\text{supp}(x))$. 

This gives further support to the phrase "dependence function", namely $\text{dep}(x, e) \{ e \}$ is the smallest set that renders $e$ dependent (again, like the fundamental circuit of a matroid). Thus, we have 1-1 mapping between all elements of $\text{sat}(x)$ and $\text{DEP}(x) = \{ \text{dep}(x, e) : e \in \text{sat}(x) \}$. 

Prof. Jeff Bilmes
EE595A/Spr 2011/Submodular Functions – Lecture 17 - May 27th, 2011
Also, for polymatroidal $f$, we saw earlier that for each $e \in \text{sat}(x) \setminus \text{supp}(x)$, the set $\text{supp}(x) + e$ is also tight. This follows since $x(\text{supp}(x) + e) = x(\text{supp}(x))$ but $e$ is dependent on $\text{supp}(x)$ so that $f(\text{supp}(x) + e) = f(\text{supp}(x))$.

Now for any point $a, b \in \text{sat}(x) \setminus \text{supp}(x)$, we have that $\text{dep}(x, a) \neq \text{dep}(x, b)$.
Also, for polymatroidal $f$, we saw earlier that for each $e \in \text{sat}(x) \setminus \text{supp}(x)$, the set $\text{supp}(x) + e$ is also tight. This follows since $x(\text{supp}(x) + e) = x(\text{supp}(x))$ but $e$ is dependent on $\text{supp}(x)$ so that $f(\text{supp}(x) + e) = f(\text{supp}(x))$.

Now for any point $a, b \in \text{sat}(x) \setminus \text{supp}(x)$, we have that $	ext{dep}(x, a) \neq \text{dep}(x, b)$

This follows, since the minimal tight set containing $a$ would never contain $b$ (and in this case, vice versa).
Also, for polymatroidal \( f \), we saw earlier that for each \( e \in \text{sat}(x) \setminus \text{supp}(x) \), the set \( \text{supp}(x) + e \) is also tight. This follows since \( x(\text{supp}(x) + e) = x(\text{supp}(x)) \) but \( e \) is dependent on \( \text{supp}(x) \) so that \( f(\text{supp}(x) + e) = f(\text{supp}(x)) \).

Now for any point \( a, b \in \text{sat}(x) \setminus \text{supp}(x) \), we have that \( \text{dep}(x, a) \neq \text{dep}(x, b) \)

- This follows, since the minimal tight set containing \( a \) would never contain \( b \) (and in this case, vice versa).
- I.e., in such case, we can have for \( a \in \text{sat}(x) \setminus \text{supp}(x) \), \( \text{dep}(x, a) = B_j + a \) for some \( j \), the smallest \( j \) such that \( f(B_j + a) = f(B_j) \), and note that \( a \notin B_j \).
Also, for polymatroidal $f$, we saw earlier that for each $e \in \text{sat}(x) \setminus \text{supp}(x)$, the set $\text{supp}(x) + e$ is also tight. This follows since $x(\text{supp}(x) + e) = x(\text{supp}(x))$ but $e$ is dependent on $\text{supp}(x)$ so that $f(\text{supp}(x) + e) = f(\text{supp}(x))$.

Now for any point $a, b \in \text{sat}(x) \setminus \text{supp}(x)$, we have that $\text{dep}(x, a) \neq \text{dep}(x, b)$

- This follows, since the minimal tight set containing $a$ would never contain $b$ (and in this case, vice versa).
- I.e., in such case, we can have for $a \in \text{sat}(x) \setminus \text{supp}(x)$, $\text{dep}(x, a) = B_j + a$ for some $j$, the smallest $j$ such that $f(B_j + a) = f(B_j)$, and note that $a \notin B_j$.

This gives further support to the phrase “dependence function”, namely $\text{dep}(x, e) \setminus \{e\} = B_j$ is the smallest set that renders $e$ dependent (again, like the fundamental circuit of a matroid).
Also, for polymatroidal \( f \), we saw earlier that for each \( e \in \text{sat}(x) \setminus \text{supp}(x) \), the set \( \text{supp}(x) + e \) is also tight. This follows since \( x(\text{supp}(x) + e) = x(\text{supp}(x)) \) but \( e \) is dependent on \( \text{supp}(x) \) so that \( f(\text{supp}(x) + e) = f(\text{supp}(x)) \).

Now for any point \( a, b \in \text{sat}(x) \setminus \text{supp}(x) \), we have that \( \text{dep}(x, a) \neq \text{dep}(x, b) \)

- This follows, since the minimal tight set containing \( a \) would never contain \( b \) (and in this case, vice versa).
- I.e., in such case, we can have for \( a \in \text{sat}(x) \setminus \text{supp}(x) \), \( \text{dep}(x, a) = B_j + a \) for some \( j \), the smallest \( j \) such that \( f(B_j + a) = f(B_j) \), and note that \( a \notin B_j \).

This gives further support to the phrase “dependence function”, namely \( \text{dep}(x, e) \setminus \{e\} = B_j \) is the smallest set that renders \( e \) dependent (again, like the fundamental circuit of a matroid).

Thus, we have 1-1 mapping between all elements of \( \text{sat}(x) \) and \( \text{DEP}(x) = \{\text{dep}(x, e) : e \in \text{sat}(x)\} \).
Therefore, the partial order on \( \text{DEP}(x) \) can be used to define a partial order on \( \text{sat}(x) \).
Therefore, the partial order on \( \text{DEP}(x) \) can be used to define a partial order on \( \text{sat}(x) \).

Now, for \( d, e \in \text{sat}(x) \), when can we have that \( \text{dep}(x, d) \subseteq \text{dep}(x, e) \)?
Therefore, the partial order on \( \text{DEP}(x) \) can be used to define a partial order on \( \text{sat}(x) \).

Now, for \( d, e \in \text{sat}(x) \), when can we have that \( \text{dep}(x, d) \subset \text{dep}(x, e) \)?

We already saw, this happens iff \( d \in \text{dep}(x, e) \).
Therefore, the partial order on $\text{DEP}(x)$ can be used to define a partial order on $\text{sat}(x)$.

Now, for $d, e \in \text{sat}(x)$, when can we have that $\text{dep}(x, d) \subset \text{dep}(x, e)$?

We already saw, this happens iff $d \in \text{dep}(x, e)$.

Thus, we can define a partial order on the elements of $\text{sat}(x)$ as follows:
Therefore, the partial order on \( \text{DEP}(x) \) can be used to define a partial order on \( \text{sat}(x) \).

Now, for \( d, e \in \text{sat}(x) \), when can we have that \( \text{dep}(x, d) \subset \text{dep}(x, e) \)?

We already saw, this happens iff \( d \in \text{dep}(x, e) \).

Thus, we can define a partial order on the elements of \( \text{sat}(x) \) as follows:

**Definition 2.1 (partial order on elements of \( \text{sat}(x) \))**

For \( d, e \in \text{sat}(x) \), we have

\[ d \preceq e \iff d \in \text{dep}(x, e) \]  \( (4) \)
Thus, we have just proven
Thus, we have just proven

**Theorem 3.1**

*If* $x \in P_f$ *is an extreme point, then* $\preceq$ *is a partial order on* $\text{sat}(x)$ *where for* $a, e \in \text{sat}(x)$, *the order* $\preceq$ *is defined by:* $a \preceq e$ *iff* $a \in \text{dep}(x, e)$. 
Thus, we have just proven

**Theorem 3.1**

*If* $x \in P_f$ *is an extreme point, then* $\preceq$ *is a partial order on* $\text{sat}(x)$ *where for* $a, e \in \text{sat}(x)$, *the order* $\preceq$ *is defined by:* $a \preceq e$ *iff* $a \in \text{dep}(x, e)$.

In fact, we have a stronger result that extreme points are characterized by this construct:
Thus, we have just proven

**Theorem 3.1**

\[ \text{If } x \in P_f \text{ is an extreme point, then } \preceq \text{ is a partial order on } \text{sat}(x) \text{ where for } a, e \in \text{sat}(x), \text{ the order } \preceq \text{ is defined by: } a \preceq e \text{ iff } a \in \text{dep}(x, e). \]

In fact, we have a stronger result that extreme points are characterized by this construct:

**Theorem 3.2**

\[ x \in P_f \text{ is an extreme point, iff } \text{supp}(x) \subseteq \text{sat}(x) \text{ and dep}(x, a) \neq \text{dep}(x, b) \text{ for every pair of distinct points } a, b \in \text{sat}(x). \]
If $f$ is strictly submodular, then the above order $\preceq$ is a total order on $\text{sat}(x)$. 
If $f$ is strictly submodular, then the above order $\leq$ is a total order on $\text{sat}(x)$.

Now our goal is to be able to, given an extreme point $x \in P_f$ characterize $\leq$, and in particular generate $\leq$ and thus characterize all orderings that generate $x$. 
If $f$ is strictly submodular, then the above order $\preceq$ is a total order on $\text{sat}(x)$.

Now our goal is to be able to, given an extreme point $x \in P_f$ characterize $\preceq$, and in particular generate $\preceq$ and thus characterize all orderings that generate $x$.

**Definition 3.3**

Given a partial order $\preceq$ and an ordered set $B = (e_1, e_2, \ldots, e_k)$, then $B$ is **compatible** with $\preceq$ if for all $i < j$ we have that $e_i \preceq e_j$. 
The partial order of extreme points

Theorem 3.4

Let \( x \) be an extreme point of \( P_f \) and \( \preceq \) be its partial order. Let \( B \subseteq E \) be an ordered set. Then \( B \) generates \( x \) using the greedy algorithm iff we have \( \text{supp}(x) \subseteq B \subseteq \text{sat}(x) \) and \( B \) is compatible with \( \preceq \).

Proof.

- Generate \( \Rightarrow \) Compatible: Let \( B \) generate \( x \)

...
The partial order of extreme points

**Theorem 3.4**

Let \( x \) be an extreme point of \( P_f \) and \( \preceq \) be its partial order. Let \( B \subseteq E \) be an ordered set. Then \( B \) generates \( x \) using the greedy algorithm iff we have \( \text{supp}(x) \subseteq B \subseteq \text{sat}(x) \) and \( B \) is compatible with \( \preceq \).

**Proof.**

- Generate \( \Rightarrow \) Compatible: Let \( B \) generate \( x \)
- Then \( \text{supp}(x) \subseteq B \).
Theorem 3.4

Let $x$ be an extreme point of $P_f$ and $\preceq$ be its partial order. Let $B \subseteq E$ be an ordered set. Then $B$ generates $x$ using the greedy algorithm iff we have $\text{supp}(x) \subseteq B \subseteq \text{sat}(x)$ and $B$ is compatible with $\preceq$.

Proof.

- Generate $\Rightarrow$ Compatible: Let $B$ generate $x$
- Then $\text{supp}(x) \subseteq B$.
- Also, since $B$ is tight, $B \in \mathcal{D}(x)$, so $B \subseteq \text{sat}(x)$.  

...
the partial order of extreme points

**Theorem 3.4**

Let $x$ be an extreme point of $P_f$ and $\preceq$ be its partial order. Let $B \subseteq E$ be an ordered set. Then $B$ generates $x$ using the greedy algorithm iff we have $\text{supp}(x) \subseteq B \subseteq \text{sat}(x)$ and $B$ is compatible with $\preceq$.

**Proof.**

- Generate $\Rightarrow$ Compatible: Let $B$ generate $x$
- Then $\text{supp}(x) \subseteq B$.
- Also, since $B$ is tight, $B \in \mathcal{D}(x)$, so $B \subseteq \text{sat}(x)$.
- Moreover, $B_j \in \mathcal{D}(x)$ (for $1 \leq j \leq |B|$), so that $\text{dep}(x, e_j) \subseteq B_j$ for $e_j$ the $j$'th element of $B$ (note $\text{dep}(x, e_j) \subseteq B_j$ if $(\text{sat}(x) \setminus \text{supp}(x)) \cap B_j = \emptyset$).
The partial order of extreme points

Theorem 3.4

Let $x$ be an extreme point of $P_f$ and $\preceq$ be its partial order. Let $B \subseteq E$ be an ordered set. Then $B$ generates $x$ using the greedy algorithm iff we have $\text{supp}(x) \subseteq B \subseteq \text{sat}(x)$ and $B$ is compatible with $\preceq$.

Proof.

- Generate $\Rightarrow$ Compatible: Let $B$ generate $x$
- Then $\text{supp}(x) \subseteq B$.
- Also, since $B$ is tight, $B \in \mathcal{D}(x)$, so $B \subseteq \text{sat}(x)$.
- Moreover, $B_j \in \mathcal{D}(x)$ (for $1 \leq j \leq |B|$), so that $\text{dep}(x, e_j) \subseteq B_j$ for $e_j$ the $j$'th element of $B$ (note $\text{dep}(x, e_j) \subseteq B_j$ if $(\text{sat}(x) \setminus \text{supp}(x)) \cap B_j = \emptyset$).
- But $g \notin B_j$ means $g \notin \text{dep}(x, e_j)$, which means $g \npreceq e_j$, meaning $B$ is compatible with $\preceq$.

...
The partial order of extreme points

**Theorem 3.4**

Let \( x \) be an extreme point of \( P_f \) and \( \preceq \) be its partial order. Let \( B \subseteq E \) be an ordered set. Then \( B \) generates \( x \) using the greedy algorithm iff we have \( \text{supp}(x) \subseteq B \subseteq \text{sat}(x) \) and \( B \) is compatible with \( \preceq \).

**Proof.**

- Conversely (Compatible \( \Rightarrow \) Generate): Suppose ordering \( B \) is compatible with \( \preceq \) and that \( \text{supp}(x) \subseteq B \subseteq \text{sat}(x) \).
The partial order of extreme points

**Theorem 3.4**

Let $x$ be an extreme point of $P_f$ and $\preceq$ be its partial order. Let $B \subseteq E$ be an ordered set. Then $B$ generates $x$ using the greedy algorithm iff we have $\text{supp}(x) \subseteq B \subseteq \text{sat}(x)$ and $B$ is compatible with $\preceq$.

**Proof.**

- Conversely (Compatible $\Rightarrow$ Generate): Suppose ordering $B$ is compatible with $\preceq$ and that $\text{supp}(x) \subseteq B \subseteq \text{sat}(x)$.

- Then for each $j$ (with $1 \leq j \leq |B|$), and for each $e \in B_j$, we have $\text{dep}(x, e) \subseteq B_j$. 

...
The partial order of extreme points

**Theorem 3.4**

Let \( x \) be an extreme point of \( P_f \) and \( \preceq \) be its partial order. Let \( B \subseteq E \) be an ordered set. Then \( B \) generates \( x \) using the greedy algorithm iff we have \( \text{supp}(x) \subseteq B \subseteq \text{sat}(x) \) and \( B \) is compatible with \( \preceq \).

**Proof.**

- Conversely (Compatible \( \Rightarrow \) Generate): Suppose ordering \( B \) is compatible with \( \preceq \) and that \( \text{supp}(x) \subseteq B \subseteq \text{sat}(x) \).

- Then for each \( j \) (with \( 1 \leq j \leq |B| \)), and for each \( e \in B_j \), we have \( \text{dep}(x, e) \subseteq B_j \).

- Thus, \( B_j \) is the union of tight sets (since each of \( \text{dep}(x, e) \) is tight), so that \( B_j \) is also tight (unions of tight sets are tight).

...
The partial order of extreme points

**Theorem 3.4**

Let $x$ be an extreme point of $P_f$ and $\preceq$ be its partial order. Let $B \subseteq E$ be an ordered set. Then $B$ generates $x$ using the greedy algorithm iff we have $\text{supp}(x) \subseteq B \subseteq \text{sat}(x)$ and $B$ is compatible with $\preceq$.

**Proof.**

- Conversely (Compatible $\Rightarrow$ Generate): Suppose ordering $B$ is compatible with $\preceq$ and that $\text{supp}(x) \subseteq B \subseteq \text{sat}(x)$.

- Then for each $j$ (with $1 \leq j \leq |B|$), and for each $e \in B_j$, we have $\text{dep}(x, e) \subseteq B_j$.

- Thus, $B_j$ is the union of tight sets (since each of $\text{dep}(x, e)$ is tight), so that $B_j$ is also tight (unions of tight sets are tight).

- Thus $B$ is tight and thus $x$ is generated by the ordering given in $B$ (by the greedy algorithm).
Extreme Points and Greedy

- We immediately get the following corollary:
We immediately get the following corollary:

**Corollary 3.5**

*If $x$ is an extreme point of $P_f$ and $B \subseteq E$ is given such that $\text{supp}(x) \subseteq B \subseteq \text{sat}(x)$, then $x$ is generated using greedy by some ordering of $B$.***
Extreme Points and Greedy

- We immediately get the following corollary:

**Corollary 3.5**

*If* \( x \) *is an extreme point of* \( P_f \) *and* \( B \subseteq E \) *is given such that* \( \text{supp}(x) \subseteq B \subseteq \text{sat}(x) \), *then* \( x \) *is generated using greedy by some ordering of* \( B \).

- this is a more satisfying way to given an extreme point show that the greedy algorithm can generate it than to resort to the polyhedral \( cv = \max(cx : x \in P_f) \) for an appropriate direction \( c \).
Extreme Points and Greedy

- We immediately get the following corollary:

**Corollary 3.5**

*If* $x$ *is an extreme point of* $P_f$ *and* $B \subseteq E$ *is given such that* $\text{supp}(x) \subseteq B \subseteq \text{sat}(x)$, *then* $x$ *is generated using greedy by some ordering of* $B$.

- This is a more satisfying way to given an extreme point show that the greedy algorithm can generate it than to resort to the polyhedral $cv = \max(cx : x \in P_f)$ for an appropriate direction $c$.
- Moreover, we can produce an efficient $O(|E|^2)$ algorithm that can produce $\preceq$ for any extreme point $x$ of $P_f$. 
Extreme Points and Greedy

- We immediately get the following corollary:

**Corollary 3.5**

*If \( x \) is an extreme point of \( P_f \) and \( B \subseteq E \) is given such that\(*
\[ \text{supp}(x) \subseteq B \subseteq \text{sat}(x), \]
then \( x \) is generated using greedy by some ordering of \( B \).*

- This is a more satisfying way to given an extreme point show that the greedy algorithm can generate it than to resort to the polyhedral

\[ cv = \max(cx : x \in P_f) \]

for an appropriate direction \( c \).

- Moreover, we can produce an efficient \( O(|E|^2) \) algorithm that can produce \( \preceq \) for any extreme point \( x \) of \( P_f \).

- The algorithm does so by, for each \( e \in \text{sat}(x) \), producing the sets

\[ \text{dep}(x, e) \]

that define the order (or otherwise terminating by saying that \( x \) is not an extreme point).
Extreme point testing and partial order generation

**input**: Vector $x \in \mathbb{R}^E$, polymatroid function $f$ on $E$.

**output**: That $x$ is not extreme point, or if it is, minimal tight sets $\text{dep}(x, e)$ for $e \in \text{sat}(x)$ thus defining $\preceq$. Moreover, $\text{dep}(x, e_j) = A_j$ for $1 \leq j \leq n$ where $n = |\text{sat}(x)|$.

1. $j \leftarrow 0$; $B \leftarrow \emptyset$
2. while true do
3.   $j \leftarrow j + 1$
4.   if $\exists e \in E \setminus B$ with $x(B + e) = f(B + e)$ then
5.     $B \leftarrow B + e$; $e_j \leftarrow e$
6.   else
7.     STOP, if $\text{supp}(x) \subseteq B$ then $x$ is extreme, otherwise not.
8.     $A_j \leftarrow B$; $k \leftarrow j - 1$
9.     while $x(A_j - e_k) = f(A_j - e_k)$ and $k > 0$ do
10.    $A_j = A_j - e_k$; $k \leftarrow k - 1$
On partial order algorithm

- Lines 4-5 just greedily add tight points, breaking ties arbitrarily.
On partial order algorithm

- Lines 4-5 just greedily add tight points, breaking ties arbitrarily.
- Lines 9-10 remove elements from \( A_j \) while retaining tightness (thus achieving \( \text{dep}(x, e_j) \)).
On partial order algorithm

- Lines 4-5 just greedily add tight points, breaking ties arbitrarily.
- Lines 9-10 remove elements from $A_j$ while retaining tightness (thus achieving $\text{dep}(x, e_j)$).
- Note, algorithm uses $f$ only to test the tightness of a set relative to a vector $x$, nothing more (i.e., line 4 could be a query on if $B + e$ is tight).
On partial order algorithm

- Lines 4-5 just greedily add tight points, breaking ties arbitrarily.
- Lines 9-10 remove elements from $A_j$ while retaining tightness (thus achieving $\text{dep}(x, e_j)$).
- Note, algorithm uses $f$ only to test the tightness of a set relative to a vector $x$, nothing more (i.e., line 4 could be a query on if $B + e$ is tight).
- We can generate all orderings consistent with a partial ordering using an algorithm by Knuth/Szwarcfiter-1974.
On partial order algorithm

- Lines 4-5 just greedily add tight points, breaking ties arbitrarily.
- Lines 9-10 remove elements from $A_j$ while retaining tightness (thus achieving $\text{dep}(x, e_j)$).
- Note, algorithm uses $f$ only to test the tightness of a set relative to a vector $x$, nothing more (i.e., line 4 could be a query on if $B + e$ is tight).
- We can generate all orderings consistent with a partial ordering using an algorithm by Knuth/Szwarcfiter-1974.
- Line 1 is $O(|E|)$, and nested lines 4 (and 9) are each $O(|E|)$, so algorithm runs in $O(|E|^2)$ doing that many function evals.
On partial order algorithm

- Lines 4-5 just greedily add tight points, breaking ties arbitrarily.
- Lines 9-10 remove elements from $A_j$ while retaining tightness (thus achieving $\text{dep}(x, e_j)$).
- Note, algorithm uses $f$ only to test the tightness of a set relative to a vector $x$, nothing more (i.e., line 4 could be a query on if $B + e$ is tight).
- We can generate all orderings consistent with a partial ordering using an algorithm by Knuth/Szwarcfiter-1974.
- Line 1 is $O(|E|)$, and nested lines 4 (and 9) are each $O(|E|)$, so algorithm runs in $O(|E|^2)$ doing that many function evals.
- Thus, extreme point testing is fundamentally computationally simpler than arbitrary membership testing (recall, to test if $x \in P_f$ in general, we needed submodular function minimization).
On partial order algorithm

- Lines 4-5 just greedily add tight points, breaking ties arbitrarily.
- Lines 9-10 remove elements from $A_j$ while retaining tightness (thus achieving $\text{dep}(x, e_j)$).
- Note, algorithm uses $f$ only to test the tightness of a set relative to a vector $x$, nothing more (i.e., line 4 could be a query on if $B + e$ is tight).
- We can generate all orderings consistent with a partial ordering using an algorithm by Knuth/Szwarcfiter-1974.
- Line 1 is $O(|E|)$, and nested lines 4 (and 9) are each $O(|E|)$, so algorithm runs in $O(|E|^2)$ doing that many function evals.
- Thus, extreme point testing is fundamentally computationally simpler than arbitrary membership testing (recall, to test if $x \in P_f$ in general, we needed submodular function minimization).
- To determine, only, if a given $x$ is extreme, we can delete lines 8-10 (having same run time).
Maximal in a tight set

**Theorem 3.6**

*Given an extreme point $x \in P_f$, with $A$ tight for $x$, and if given order $\preceq$ element $e \in A$ is maximal, then $A - e$ is also tight.*

**Proof.**

- If $e$ is maximal in $A$ w.r.t. $\preceq$, then there exists no $a \in A \setminus \{e\}$, such that $e \in \text{dep}(x, a)$.
Maximal in a tight set

**Theorem 3.6**

*Given an extreme point* \( x \in P_f \), *with* \( A \) *tight for* \( x \), *and if given order* \( \preceq \) *element* \( e \in A \) *is maximal, then* \( A - e \) *is also tight.*

**Proof.**

- If \( e \) is maximal in \( A \) w.r.t. \( \preceq \), then there exists no \( a \in A \setminus \{ e \} \), such that \( e \in \text{dep}(x, a) \).
- Thus, \( \text{dep}(x, a) \subseteq A \setminus \{ e \} \) for all \( a \in A \setminus \{ e \} \).
Maximal in a tight set

**Theorem 3.6**

Given an extreme point $x \in P_f$, with $A$ tight for $x$, and if given order $\preceq$ element $e \in A$ is maximal, then $A - e$ is also tight.

**Proof.**

- If $e$ is maximal in $A$ w.r.t. $\preceq$, then there exists no $a \in A \setminus \{e\}$, such that $e \in \text{dep}(x, a)$.
- Thus, $\text{dep}(x, a) \subseteq A \setminus \{e\}$ for all $a \in A \setminus \{e\}$.
- Now, since $\text{dep}(x, a)$ is the smallest $x$-tight set containing $a$ and $\text{dep}(x, a) \subseteq A \setminus \{e\}$, we have

$$
\bigcup_{a \in A \setminus \{e\}} \text{dep}(x, a) = A \setminus \{e\}
$$

(5)

...
Maximal in a tight set

**Theorem 3.6**

*Given an extreme point $x \in P_f$, with $A$ tight for $x$, and if given order $\preceq$ element $e \in A$ is maximal, then $A - e$ is also tight.*

**Proof.**

- If $e$ is maximal in $A$ w.r.t. $\preceq$, then there exists no $a \in A \setminus \{e\}$, such that $e \in \text{dep}(x, a)$.

- Thus, $\text{dep}(x, a) \subseteq A \setminus \{e\}$ for all $a \in A \setminus \{e\}$.

- Now, since $\text{dep}(x, a)$ is the smallest $x$-tight set containing $a$ and $\text{dep}(x, a) \subseteq A \setminus \{e\}$, we have

$$
\bigcup_{a \in A \setminus \{e\}} \text{dep}(x, a) = A \setminus \{e\}
$$

(5)

- Since the union (and intersection) of tight sets is tight, we have that $A \setminus \{e\}$ is therefore also tight.
Maximal in a tight set

**Theorem 3.6**

*Given an extreme point* \( x \in P_f \), *with* \( A \) *tight for* \( x \), *and if given order* \( \leq \) *element* \( e \in A \) *is maximal, then* \( A - e \) *is also tight.*

We also have

**Corollary 3.7**

*For all* \( e \in \text{sat}(x) \), *we have that* \( \text{dep}(x, e) \setminus e \) *is also tight.*

**Proof.**
Maximal in a tight set

Theorem 3.6

Given an extreme point $x \in P_f$, with $A$ tight for $x$, and if given order $\preceq$ element $e \in A$ is maximal, then $A - e$ is also tight.

We also have

Corollary 3.7

For all $e \in \text{sat}(x)$, we have that $\text{dep}(x, e) \setminus e$ is also tight.

Proof.

- $\text{dep}(x, e)$ is tight, and recall that there is some ordered set $B_j^e$ with $\text{dep}(x, e) = B_j^e$ whose's last ($j$'th) item is $e$. 

Maximal in a tight set

**Theorem 3.6**

*Given an extreme point* $x \in P_f$, *with* $A$ *tight for* $x$, *and if given order* $\preceq$ *element* $e \in A$ *is maximal, then* $A - e$ *is also tight.*

We also have

**Corollary 3.7**

*For all* $e \in \text{sat}(x)$, *we have that* $\text{dep}(x, e) \setminus e$ *is also tight.*

**Proof.**

- $\text{dep}(x, e)$ *is tight, and recall that there is some ordered set* $B^e_j$ *with* $\text{dep}(x, e) = B^e_j$ *whose’s last* $(j’th)$ *item is* $e$.

- This theorem and corollary allow us to prove that the above algorithm gives us not only the minum min sets containing $e$ but the minimum tight sets with $e$, *i.e.,* $\text{dep}(x, e)$. 
Maximal in a tight set

**Theorem 3.6**

Given an extreme point \( x \in P_f \), with \( A \) tight for \( x \), and if given order \( \leq \) element \( e \in A \) is maximal, then \( A - e \) is also tight.

We also have

**Corollary 3.7**

For all \( e \in \text{sat}(x) \), we have that \( \text{dep}(x, e) \setminus e \) is also tight.

**Proof.**

- \( \text{dep}(x, e) \) is tight, and recall that there is some ordered set \( B_j^e \) with \( \text{dep}(x, e) = B_j^e \) whose’s last \((j’th)\) item is \( e \).

- This theorem and corollary allow us to prove that the above algorithm gives us not only the minum min sets containing \( e \) but the minimum tight sets with \( e \), i.e., \( \text{dep}(x, e) \).
Maximal in a tight set

- Also, for any $x \in P_f$, and $\forall e \in \text{sat}(x)$, we have that

$$\text{dep}(x, e) \setminus \{e\} \subseteq \text{supp}(x)$$  \hspace{1cm} (6)

- This follows since suppose $\exists e' \in \text{dep}(x, e) \setminus \{e\}$ such that $x(e') = 0$.

- Then, since $f(e') > 0$, in such case $\text{dep}(x, e)$ wouldn’t be minimally $e$-containing tight, since we’d have $x(\text{dep}(x, e) \setminus \{e'\}) = x(\text{dep}(x, e)) = f(\text{dep}(x, e))$. 

On Greedy, and linear programming max

Theorem 3.8

Let $y \in P_f$ be an extreme point, and let $\preceq$ be the partial order of $y$. Let $c \in \mathbb{R}^E$. Then, $y$ is the solution in:

$$c^T y = \max \{ c^T x : x \in P_f \} \quad (7)$$

iff the following three conditions hold:

1. $c(e) \geq 0$ for every $e \in \text{supp}(y)$
2. $c(e) \leq 0$ for every $e \in E \setminus \text{sat}(y)$, and
3. For $d, e \in \text{sat}(y)$ and $d \preceq e$ imply that $c(d) \geq c(e)$. 
Another revealing theorem

**Theorem 3.9**

Let $f$ be a polymatroid function and suppose that $E$ can be partitioned into $(E_1, E_2, \ldots, E_k)$ such that $f(A) = \sum_{i=1}^{k} f(A \cap E_i)$ for all $A \subseteq E$, and $k$ is maximum. Then the base polytope $B_f = \{x \in P_f : x(E) = f(E)\}$ (the $E$-tight subset of $P_f$) has dimension $|E| - k$. 
Another revealing theorem

Theorem 3.9

Let $f$ be a polymatroid function and suppose that $E$ can be partitioned into $(E_1, E_2, \ldots, E_k)$ such that $f(A) = \sum_{i=1}^{k} f(A \cap E_i)$ for all $A \subseteq E$, and $k$ is maximum. Then the base polytope $B_f = \{x \in P_f : x(E) = f(E)\}$ (the $E$-tight subset of $P_f$) has dimension $|E| - k$.

Thus, “independence” between disjoint $A$ and $B$ (leading to a rectangular projection of the polymatroid polytope) reduces the dimension of the base polytope, as expected.
Another revealing theorem

**Theorem 3.9**

Let $f$ be a polymatroid function and suppose that $E$ can be partitioned into $(E_1, E_2, \ldots, E_k)$ such that $f(A) = \sum_{i=1}^{k} f(A \cap E_i)$ for all $A \subseteq E$, and $k$ is maximum. Then the base polytope $B_f = \{x \in P_f : x(E) = f(E)\}$ (the $E$-tight subset of $P_f$) has dimension $|E| - k$.

- Thus, “independence” between disjoint $A$ and $B$ (leading to a rectangular projection of the polymatroid polytope) reduces the dimension of the base polytope, as expected.
- Thus, any point $x \in B_f$ is a convex combination of at most $|E| - k + 1$ vertices of $B_f$. 

Theorem 3.9

Let \( f \) be a polymatroid function and suppose that \( E \) can be partitioned into \((E_1, E_2, \ldots, E_k)\) such that \( f(A) = \sum_{i=1}^{k} f(A \cap E_i) \) for all \( A \subseteq E \), and \( k \) is maximum. Then the base polytope \( B_f = \{ x \in P_f : x(E) = f(E) \} \) (the \( E \)-tight subset of \( P_f \)) has dimension \( |E| - k \).

- Thus, “independence” between disjoint \( A \) and \( B \) (leading to a rectangular projection of the polymatroid polytope) reduces the dimension of the base polytope, as expected.
- Thus, any point \( x \in B_f \) is a convex combination of at most \( |E| - k + 1 \) vertices of \( B_f \).
- And if \( f \) does not have such independence, dimension of \( B_f \) is \( |E| - 1 \) and any point \( x \in B_f \) is a convex combination of at most \( |E| \) vertices of \( B_f \).
Another revealing theorem

**Theorem 3.9**

Let $f$ be a polymatroid function and suppose that $E$ can be partitioned into $(E_1, E_2, \ldots, E_k)$ such that $f(A) = \sum_{i=1}^{k} f(A \cap E_i)$ for all $A \subseteq E$, and $k$ is maximum. Then the base polytope $B_f = \{ x \in P_f : x(E) = f(E) \}$ (the $E$-tight subset of $P_f$) has dimension $|E| - k$.

- Example $f$ with independence between $A = \{e_2, e_3\}$ and $B = \{e_1\}$, i.e., $e_1 \perp \{e_2, e_3\}$, with $B_f$ marked in green.
Given polymatroid function $f$, the base polytope $B_f = \{ x \in \mathbb{R}_+^E : x(A) \leq f(A) \ \forall A \subseteq E, \text{ and } x(E) = f(E) \}$ always exists.
Base polytope existence

- Given polymatroid function $f$, the base polytope $B_f = \{ x \in \mathbb{R}^E_+ : x(A) \leq f(A) \ \forall A \subseteq E, \text{ and } x(E) = f(E) \}$ always exists.

- Consider any order of $E$ and generate a vector $x$ by this order (i.e., $x(e_1) = f(\{e_1\})$, $x(e_2) = f(\{e_1, e_2\}) - f(\{e_1\})$, and so on).
Given polymatroid function $f$, the base polytope $B_f = \{ x \in \mathbb{R}^E_+ : x(A) \leq f(A) \ \forall A \subseteq E, \text{ and } x(E) = f(E) \}$ always exists.

Consider any order of $E$ and generate a vector $x$ by this order (i.e., $x(e_1) = f(\{e_1\})$, $x(e_2) = f(\{e_1, e_2\}) - f(\{e_1\})$, and so on).

From past lectures, we now know that:
Base polytope existence

- Given polymatroid function $f$, the base polytope
  
  $$B_f = \{ x \in \mathbb{R}^E_+ : x(A) \leq f(A) \quad \forall A \subseteq E, \text{ and } x(E) = f(E) \}$$

  always exists.

- Consider any order of $E$ and generate a vector $x$ by this order (i.e.,
  
  $$x(e_1) = f(\{e_1\}), \quad x(e_2) = f(\{e_1, e_2\}) - f(\{e_1\}), \text{ and so on}.$$  

- From past lectures, we now know that:
  
  (1) $x \in P_f$
Base polytope existence

Given polymatroid function $f$, the base polytope

$$B_f = \left\{ x \in \mathbb{R}^E_+ : x(A) \leq f(A) \ \forall A \subseteq E, \text{ and } x(E) = f(E) \right\}$$
always exists.

Consider any order of $E$ and generate a vector $x$ by this order (i.e.,

$$x(e_1) = f(\{e_1\}), \ x(e_2) = f(\{e_1, e_2\}) - f(\{e_1\}), \text{ and so on).}$$

From past lectures, we now know that:

1. $x \in P_f$
2. $x$ is an extreme point in $P_f$
Base polytope existence

- Given polymatroid function $f$, the base polytope $B_f = \{ x \in \mathbb{R}^E_+ : x(A) \leq f(A) \forall A \subseteq E, \text{ and } x(E) = f(E) \}$ always exists.

- Consider any order of $E$ and generate a vector $x$ by this order (i.e., $x(e_1) = f(\{e_1\}), x(e_2) = f(\{e_1, e_2\}) - f(\{e_1\})$, and so on).

- From past lectures, we now know that:
  
  1. $x \in P_f$
  2. $x$ is an extreme point in $P_f$
  3. Since $x$ is generated using an ordering of all of $E$, we have that $x(E) = f(E)$. 
Given polymatroid function $f$, the base polytope $B_f = \{ x \in \mathbb{R}^E_+ : x(A) \leq f(A) \forall A \subseteq E, \text{ and } x(E) = f(E) \}$ always exists.

Consider any order of $E$ and generate a vector $x$ by this order (i.e., $x(e_1) = f(\{e_1\}), x(e_2) = f(\{e_1, e_2\}) - f(\{e_1\})$, and so on).

From past lectures, we now know that:

1. $x \in P_f$
2. $x$ is an extreme point in $P_f$
3. Since $x$ is generated using an ordering of all of $E$, we have that $x(E) = f(E)$.

Thus $x \in B_f$, and $B_f$ is never empty.
Base polytope existence

- Given polymatroid function $f$, the base polytope $B_f = \{ x \in \mathbb{R}^E_+ : x(A) \leq f(A) \ \forall A \subseteq E, \ \text{and} \ x(E) = f(E) \}$ always exists.

- Consider any order of $E$ and generate a vector $x$ by this order (i.e., $x(e_1) = f(\{e_1\})$, $x(e_2) = f(\{e_1, e_2\}) - f(\{e_1\})$, and so on).

- From past lectures, we now know that:
  
  1. $x \in P_f$
  2. $x$ is an extreme point in $P_f$
  3. Since $x$ is generated using an ordering of all of $E$, we have that $x(E) = f(E)$.

- Thus $x \in B_f$, and $B_f$ is never empty.

- Moreover, in this case, $x$ is a vertex of $B_f$ since it is extremal.
Now, for any $A \subseteq E$, we can generate a particular point in $B_f$.
Base polytope property

- Now, for any $A \subseteq E$, we can generate a particular point in $B_f$
- That is, choose the ordering of $E = (e_1, e_2, \ldots, e_n)$ where $n = |E|$, and where $E_i = (e_1, e_2, \ldots, e_i)$, so that we have $E_k = A$ with $k = |A|$.
Base polytope property

- Now, for any $A \subseteq E$, we can generate a particular point in $B_f$
- That is, choose the ordering of $E = (e_1, e_2, \ldots, e_n)$ where $n = |E|$, and where $E_i = (e_1, e_2, \ldots, e_i)$, so that we have $E_k = A$ with $k = |A|$.
- Then, we have generated a point $x$ (a vertex, no less) in $B_f$ such that $x(A) = f(A)$. 

In words, $B_f$ intersects all "multi-axis orthogonal" subsets of $R^E_+$. 
Base polytope property

- Now, for any $A \subseteq E$, we can generate a particular point in $B_f$
- That is, choose the ordering of $E = (e_1, e_2, \ldots, e_n)$ where $n = |E|$, and where $E_i = (e_1, e_2, \ldots, e_i)$, so that we have $E_k = A$ with $k = |A|$.
- Then, we have generated a point $x$ (a vertex, no less) in $B_f$ such that $x(A) = f(A)$.
- Thus, for any $A$, we have

$$B_f \cap \left\{ x \in \mathbb{R}_+^E : x(A) = f(A) \right\} \neq \emptyset$$

(8)
Base polytope property

- Now, for any $A \subseteq E$, we can generate a particular point in $B_f$
- That is, choose the ordering of $E = (e_1, e_2, \ldots, e_n)$ where $n = |E|$, and where $E_i = (e_1, e_2, \ldots, e_i)$, so that we have $E_k = A$ with $k = |A|$.
- Then, we have generated a point $x$ (a vertex, no less) in $B_f$ such that $x(A) = f(A)$.
- Thus, for any $A$, we have

$$B_f \cap \left\{ x \in \mathbb{R}_+^E : x(A) = f(A) \right\} \neq \emptyset \quad (8)$$

- In words, $B_f$ intersects all “multi-axis orthogonal” subsets of $\mathbb{R}_+^E$. 

Polytope example 1

Observe: $P_f$ (at two views):

Is this a polymatroidal polytope?

No, $B_f$ doesn’t intersect sets of the form $\{x: x(e) = f(e)\}$ for $e \in E$. This was generated using function $g(0) = 0$, $g(1) = 3$, $g(2) = 4$, and $g(3) = 5$. Then $f(S) = g(|S|)$ is not submodular since (e.g.) $f(\{e_1, e_3\}) + f(\{e_1, e_2\}) = 4 + 4 = 8$ but $f(\{e_1, e_2, e_3\}) + f(\{e_1\}) = 5 + 3 = 8$.5.
Polytope example 1

- Observe: $P_f$ (at two views):

- Is this a polymatroidal polytope?

- Is this a polymatroidal polytope?
Polytope example 1

- Observe: $P_f$ (at two views):

Is this a polymatroidal polytope?
- No, "$B_f$" doesn't intersect sets of the form $\{x : x(e) = f(e)\}$ for $e \in E$. 

Prof. Jeff Bilmes
Polytope example 1

- Observe: $P_f$ (at two views):

Is this a polymatroidal polytope?
No, “$B_f$” doesn’t intersect sets of the form \{x : x(e) = f(e)\} for $e \in E$.

This was generated using function $g(0) = 0$, $g(1) = 3$, $g(2) = 4$, and $g(3) = 5.5$. Then $f(S) = g(|S|)$ is not submodular since (e.g.)
\[
f(\{e_1, e_3\}) + f(\{e_1, e_2\}) = 4 + 4 = 8 \text{ but}
\]
Polytope example 2

- Observe: $P_f$ (at two views):

Is this a polymatroidal polytope?

No, $B_f$ (which would be a single point in this case) doesn't intersect sets of the form

\{x : x(e) = f(e)\}

for $e \in E$.

This was generated using function $g(0) = 0$, $g(1) = 1$, $g(2) = 1.8$, and $g(3) = 3$. Then $f(S) = g(|S|)$ is not submodular since (e.g.)

$f(\{e_1, e_3\}) + f(\{e_1, e_2\}) = 1.8 + 1.8 = 3.6$ but $f(\{e_1, e_2, e_3\}) + f(\{e_1\}) = 3 + 1 = 4.$
Polytope example 2

- Observe: $P_f$ (at two views):

- Is this a polymatroidal polytope?
Polytope example 2

- Observe: \( P_f \) (at two views):

- Is this a polymatroidal polytope?
- No, \( B_f \) (which would be a single point in this case) doesn’t intersect sets of the form \( \{ x : x(e) = f(e) \} \) for \( e \in E \).
Polytope example 2

- Observe: $P_f$ (at two views):

Is this a polymatroidal polytope?

No, “$B_f$” (which would be a single point in this case) doesn’t intersect sets of the form \( \{x : x(e) = f(e)\} \) for \( e \in E \).

This was generated using function \( g(0) = 0, g(1) = 1, g(2) = 1.8, \) and \( g(3) = 3 \). Then \( f(S) = g(|S|) \) is not submodular since (e.g.)

\[
\begin{align*}
  f(\{e_1, e_3\}) + f(\{e_1, e_2\}) &= 1.8 + 1.8 = 3.6 \\
  f(\{e_1, e_2, e_3\}) + f(\{e_1\}) &= 3 + 1 = 4 .
\end{align*}
\]
First, given any submodular function $g$, construct modular function $m : E \rightarrow \mathbb{R}$ such that $m(e) = g(E \setminus \{e\}) - g(E)$.

Note that

$$m(e) = g(E \setminus \{e\}) - g(E)$$
$$= -[g(E) - g(E \setminus e)]$$
$$= -[\text{gain of adding } e \text{ to } E \setminus e]$$
$$= -[\text{smallest possible gain/value of } e \text{ in any context} ]$$

The last equality follows from submodularity.
First, given any submodular function $g$, construct modular function $m : E \rightarrow \mathbb{R}$ such that $m(e) = g(E \setminus \{e\}) - g(E)$.

Then construct a new function $f : 2^E \rightarrow \mathbb{R}_+$ as

$$f(A) = g(A) + m(A) - g(\emptyset)$$  (13)
First, given any submodular function $g$, construct modular function $m : E \rightarrow \mathbb{R}$ such that $m(e) = g(E \setminus \{e\}) - g(E)$.

Then construct a new function $f : 2^E \rightarrow \mathbb{R}_+$ as

$$f(A) = g(A) + m(A) - g(\emptyset) \quad (13)$$

Then $f(\emptyset) = 0$, so $f$ is normalized.
First, given any submodular function $g$, construct modular function $m : E \rightarrow \mathbb{R}$ such that $m(e) = g(E \setminus \{e\}) - g(E)$.

Then construct a new function $f : 2^E \rightarrow \mathbb{R}_+$ as

$$f(A) = g(A) + m(A) - g(\emptyset)$$  \hspace{1cm} (13)

Then $f(\emptyset) = 0$, so $f$ is normalized.

Also, $f$ is monotone non-decreasing (and thus non-negative) and submodular. It is submodular since sum of submodular and modular. Monotone non-decreasing follows since for $v \notin B$,

$$f(B + v) - f(B) = g(B + v) - g(B) + m(v)$$  \hspace{1cm} (14)

$$= g(B + v) - g(B) + g(E - v) - g(E)$$  \hspace{1cm} (15)

$$\geq 0$$  \hspace{1cm} (16)

since, by submodularity, $g(B + v) - g(B) \geq g(E) - g(E - v)$. 
SFM for arbitrary submodular $g$ (from lecture 11)

- Also, if we wish to minimize arbitrary submodular $g$, then given $f(A) = g(A) + m(A) - g(\emptyset)$, we can just minimize $f - m$ since $g(\emptyset)$ is a constant.
Also, if we wish to minimize arbitrary submodular $g$, then given $f(A) = g(A) + m(A) - g(\emptyset)$, we can just minimize $f - m$ since $g(\emptyset)$ is a constant.

So now we have a difference of a polymatroid function $f$ and a modular function $m$. 
Dealing with $m \in \mathbb{R}^E_+$

So now we reduced the problem of SFM to that of minimizing a difference between a polymatroid function $f$ and a modular function $m$ (i.e., $\min_{A \subseteq E} f(A) - m(A)$).
Dealing with $m \in \mathbb{R}^E_+$

- So now we reduced the problem of SFM to that of minimizing a difference between a polymatroid function $f$ and a modular function $m$ (i.e., $\min_{A \subseteq E} f(A) - m(A)$).

- Is $m \in \mathbb{R}^E_+$?
Dealing with $m \in \mathbb{R}_+^E$

- So now we reduced the problem of SFM to that of minimizing a difference between a polymatroid function $f$ and a modular function $m$ (i.e., $\min_{A \subseteq E} f(A) - m(A)$).

- Is $m \in \mathbb{R}_+^E$?

No, not in general, but for any $e$ such that $m(e) < 0$, $e$ can’t be a minimizer of $f - m$ since, assuming that $A$ minimizes $f(A) - m(A)$ and $e \in A$ is such that $m(e) < 0$, then we have that $f(A') - m(A') < f(A) - m(A)$ where $A' = A \setminus \{e\}$.
Dealing with $m \in \mathbb{R}_+^E$

- So now we reduced the problem of SFM to that of minimizing a difference between a polymatroid function $f$ and a modular function $m$ (i.e., $\min_{A \subseteq E} f(A) - m(A)$).

- Is $m \in \mathbb{R}_+^E$?

- No, not in general, but for any $e$ such that $m(e) < 0$, $e$ can’t be a minimizer of $f - m$ since, assuming that $A$ minimizes $f(A) - m(A)$ and $e \in A$ is such that $m(e) < 0$, then we have that $f(A') - m(A') < f(A) - m(A)$ where $A' = A \setminus \{e\}$.

- This follows since $f$ is monotone non-decreasing, and $m(A) = m(A') + m(e)$, so $m(A') > m(A)$.
Dealing with $m \in \mathbb{R}_+^E$

- So now we reduced the problem of SFM to that of minimizing a difference between a polymatroid function $f$ and a modular function $m$ (i.e., $\min_{A \subseteq E} f(A) - m(A)$).
- Is $m \in \mathbb{R}_+^E$?
- No, not in general, but for any $e$ such that $m(e) < 0$, $e$ can’t be a minimizer of $f - m$ since, assuming that $A$ minimizes $f(A) - m(A)$ and $e \in A$ is such that $m(e) < 0$, then we have that $f(A') - m(A') < f(A) - m(A)$ where $A' = A \setminus \{e\}$.
- This follows since $f$ is monotone non-decreasing, and $m(A) = m(A') + m(e)$, so $m(A') > m(A)$.
- So we “throw away” any $e$ s.t. $m(e) < 0$. We get a new function on $E' = E \setminus M$ where $M = \{e : m(e) < 0\}$, and define new function $f' : 2^{E'} \rightarrow \mathbb{R}_+$ with $f'(A) = f(A)$ for $A \subseteq E'$. This deals with (2) above.
SFM on arbitrary submodular $g$: transformation

- Given any arbitrary submodular function $g$ with the goal of finding $A^* \in \text{argmin}_{A \subseteq E} g(A)$
- We reduce this to:
  
  $$A^* \in \text{argmin}_{A \subseteq E'} \left(f(A) - m(A)\right)$$  

  \hspace{1cm} (17)

  where

  - $f$ is a polymatroid function on $2^{E'}$
  - $m$ is a modular function on $2^{E'}$ with $m \in \mathbb{R}^{E'}_+.$
  - $E' \subseteq E.$

- In the sequel, we assume this form, with ground set $E$.
- Moreover, we may assume that $P_f$ is a polymatroidal polytope, with $P_f \subset \mathbb{R}^E_+.$
A characterization of the optimality of the SFM problem

- We proved in Lecture 7 (and again, in Lecture 11) the following theorem:
A characterization of the optimality of the SFM problem

- We proved in Lecture 7 (and again, in Lecture 11) the following theorem:

**Theorem 4.1**

*Let $f$ be a polymatroid function defined on subsets of $E$. For any $x \in \mathbb{R}_+^E$, then*

$$\max \left( y(E) : y \leq x, y \in P_f \right) = \min \left( f(A) + x(E \setminus A) : A \subseteq E \right)$$

(18)

This can act as a certificate of optimality for any submodular function minimization problem on $E$, even if $g$ is not polymatroidal. We need only find a feasible $y$ on the max (left) side, and an $A^*$ on the min (right) side that achieves equality, then $A^*$ is a SFM solution in $A^* \in \arg\min_{A \subseteq E} g(A)$ where $x$ is the aforementioned modular function, and $f(A) = g(A) + m(A) - g(\emptyset)$. 

Prof. Jeff Bilmes
EE595A/Spr 2011/Submodular Functions – Lecture 17 - May 27th, 2011
A characterization of the optimality of the SFM problem

- We proved in Lecture 7 (and again, in Lecture 11) the following theorem:

**Theorem 4.1**

Let $f$ be a polymatroid function defined on subsets of $E$. For any $x \in \mathbb{R}^E_+$, then

$$\max \{ y(E) : y \leq x, y \in P_f \} = \min \{ f(A) + x(E \setminus A) : A \subseteq E \} \quad (18)$$

- Thus, this can act as a certificate of optimality for any submodular function minimization problem on $g$ even if $g$ is not polymatroidal.
A characterization of the optimality of the SFM problem

- We proved in Lecture 7 (and again, in Lecture 11) the following theorem:

**Theorem 4.1**

Let $f$ be a polymatroid function defined on subsets of $E$. For any $x \in \mathbb{R}^E_+$, then

$$\max (y(E) : y \leq x, y \in P_f) = \min (f(A) + x(E \setminus A) : A \subseteq E)$$

Thus, this can act as a certificate of optimality for any submodular function minimization problem on $g$ even if $g$ is not polymatroidal.

- We need only find a feasible $y$ on the max (left) side, and an $A^*$ on the min (right) side that achieves equality, then $A^*$ is a SFM solution in $A^* \in \arg\min_{A \subseteq E} g(A)$ where $x$ is the aforementioned modular function, and $f(A) = g(A) + m(A) - g(\emptyset)$. 

---

Prof. Jeff Bilmes  
EE595A/Spr 2011/Submodular Functions – Lecture 17 - May 27th, 2011  page 29
Maximizing $y$

- The nature of SFM will be very similar to the Edmonds's matroid partition problem (recall, asking if $E$ can be partitioned into $\{I_i\}$ each independent in a matroid $M_i$) and the core algorithm is very similar.
Maximizing $y$

- The nature of SFM will be very similar to the Edmonds’s matroid partition problem (recall, asking if $E$ can be partitioned into $\{I_i\}$ each independent in a matroid $M_i$) and the core algorithm is very similar.

- Now, from convex polytope theory, any $x \in P_f$ can be represented as a convex combination of at most $|E| + 1$ extreme points of $P_f$ (each of which may be generated by greedy).
Maximizing $y$

- The nature of SFM will be very similar to the Edmonds's matroid partition problem (recall, asking if $E$ can be partitioned into $\{I_i\}$ each independent in a matroid $M_i$) and the core algorithm is very similar.

- Now, from convex polytope theory, any $x \in P_f$ can be represented as a convex combination of at most $|E| + 1$ extreme points of $P_f$ (each of which may be generated by greedy).

- We keep a feasible solution to the max version of the problem as a convex combination of such extreme points.
Maximizing $y$

- The nature of SFM will be very similar to the Edmonds's matroid partition problem (recall, asking if $E$ can be partitioned into $\{I_i\}$ each independent in a matroid $M_i$) and the core algorithm is very similar.

- Now, from convex polytope theory, any $x \in P_f$ can be represented as a convex combination of at most $|E| + 1$ extreme points of $P_f$ (each of which may be generated by greedy).

- We keep a feasible solution to the max version of the problem as a convex combination of such extreme points.

- That is, let $I$ be an index set, and $x^{(i)}$ be an extreme point of $P_f$ for $i \in I$. We then keep $y$ as

$$y = \sum_{i \in I} \lambda_i x^{(i)} \quad (19)$$

where $\lambda_i$ are convex coefficients.
Maximizing $y$

- The nature of SFM will be very similar to the Edmonds’s matroid partition problem (recall, asking if $E$ can be partitioned into $\{I_i\}$ each independent in a matroid $M_i$) and the core algorithm is very similar.

- Now, from convex polytope theory, any $x \in P_f$ can be represented as a convex combination of at most $|E| + 1$ extreme points of $P_f$ (each of which may be generated by greedy).

- We keep a feasible solution to the max version of the problem as a convex combination of such extreme points.

- That is, let $I$ be an index set, and $x^{(i)}$ be an extreme point of $P_f$ for $i \in I$. We then keep $y$ as

$$y = \sum_{i \in I} \lambda_i x^{(i)}$$

(19)

where $\lambda_i$ are convex coefficients.

- At each step of the algorithm, we either find a larger $y$, or demonstrate $y$’s optimality by finding a minimizing $A$. 

Maximizing $y$

- The nature of SFM will be very similar to the Edmonds’s matroid partition problem (recall, asking if $E$ can be partitioned into $\{I_i\}$ each independent in a matroid $M_i$) and the core algorithm is very similar.

- Now, from convex polytope theory, any $x \in P_f$ can be represented as a convex combination of at most $|E| + 1$ extreme points of $P_f$ (each of which may be generated by greedy).

- We keep a feasible solution to the max version of the problem as a convex combination of such extreme points.

- That is, let $I$ be an index set, and $x^{(i)}$ be an extreme point of $P_f$ for $i \in I$. We then keep $y$ as

$$y = \sum_{i \in I} \lambda_i x^{(i)}$$

where $\lambda_i$ are convex coefficients.

- At each step of the algorithm, we either find a larger $y$, or demonstrate $y$’s optimality by finding a minimizing $A$.

- Start with $y = 0$, $I = \{1\}$, $\lambda_1 = 1$, and $\nu^{(1)} = 0$. 
From vertex to vertex

- We will need to move from one extreme point to another (adjacent) extreme point, and will use an augmenting path like approach to do so.
- How do we characterize such adjacent extreme points?
From vertex to vertex

Theorem 4.2

Let $x$ be an extreme point of $P_f$, and let $\preceq$ be its partial order. Then, each of the following three operations will yield a new extreme point $w$:

(a) Let $a, b \in E$ and $a$ cover $b$ relative to $\preceq$. Let $w = x + \alpha \mathbf{1}_a - \alpha \mathbf{1}_b$ with $\alpha = f(\text{dep}(x, a) - b) - x(\text{dep}(x, a) - b)$.
From vertex to vertex

**Theorem 4.2**

Let $x$ be an extreme point of $P_f$, and let $\preceq$ be its partial order. Then, each of the following three operations will yield a new extreme point $w$:

(a) Let $a, b \in E$ and $a$ cover $b$ relative to $\preceq$. Let $w = x + \alpha 1_a - \alpha 1_b$ with $
abla = f(\text{dep}(x, a) - b) - x(\text{dep}(x, a) - b)$.

(b) Let $a \in E \setminus \text{sat}(x)$, and let $w = x + \alpha 1_a$ where $\alpha = f(\text{sat}(x) + a) - f(\text{sat}(x))$. 
From vertex to vertex

Theorem 4.2

Let $x$ be an extreme point of $P_f$, and let $\leq$ be its partial order. Then, each of the following three operations will yield a new extreme point $w$:

(a) Let $a, b \in E$ and $a$ cover $b$ relative to $\leq$. Let $w = x + \alpha \mathbf{1}_a - \alpha \mathbf{1}_b$ with $\alpha = f(\text{dep}(x, a) - b) - x(\text{dep}(x, a) - b)$.

(b) Let $a \in E \setminus \text{sat}(x)$, and let $w = x + \alpha \mathbf{1}_a$ where $\alpha = f(\text{sat}(x) + a) - f(\text{sat}(x))$.

(c) Let $a \in \text{supp}(x)$ be maximal (w.r.t. $\leq$), and let $w = x - x(a) \mathbf{1}_a$. 
Sources for Today’s Lecture