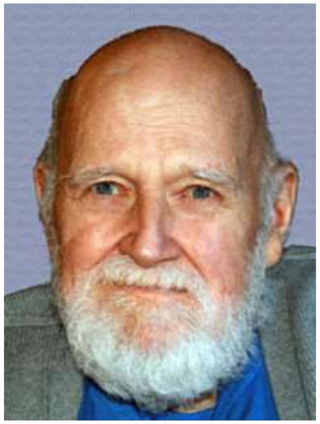
Lecture 7: Min-Cost Transshipment

Outline

- Min-Cost Max-Flow: shortest-path augmenting
- Min-Cost Uncapacitated TS: wide shortest-path augmenting
- Min-Cost Circulation: circuit canceling

Assumption: nonnegative edge capacities

1. Min-Cost Max-Flow



Generic method for min-cost max-flow

Assumption: no negative circuits (so that 0 is an extreme flow)

Idea:

- Maintain an extreme flow f (initially 0)
- Progress toward max-flow (excess feasibility)

```
Ford-Fulkerson [1958]:
    f ← 0;
    while there is an s-t path in D_f
        find a shortest s-t path P in D_f
        f ← f ⊕ P;
    return f
```

Invariant: f is an extreme flow

Distance-based potentials

Remark: same as in the **Hungarian algorithm** for Max-Weight BM

Def. $p_f(v) :=$ distance from s to v in D_f, for each $v \in V$.

Claim. p_f is a potential for $D_{f \oplus P}$

Pf. For each $a = (u, v) \in A_{f \oplus P} \setminus A_f \subseteq P^{-1}$, $(v, u) \in P$ hence

$$p_f(v) = p_f(u) - \ell(v, u) = p_f(u) + \ell(u, v)$$

Implementation

```
MinCost-MF(G)

    Initialization of f=0,p
    repeat
        reweight all arcs in D_f using p
        apply Dijkstra's algorithm to compute P and p
        if P is not found return f
        else f ← f ⊕ P
```

Initialization of p: Dijkstra's algorithm if nonnegative edge prices or acyclic, or Bellman-Ford algorithm in general

Excluding the initialization, $O(\varphi)$ Dijkstra's SP-computations, where φ is the max-flow value

2. Min-Cost Uncapacitated TS



Min-Cost Uncapacitated b-TS

Assumptions:

- no negative circuits (otherwise, no finite solution)
- D is strongly connected.
  - if necessary, add artificial arcs (s, v) and (v, s) for each v and assigning a large price. No such arc would appear in a minimum cost solution.

Idea:

- Maintain an extreme pseudoflow f (initially 0)
- Progress toward excess feasibility

Excess gap and feasibility gap

f: a pseudoflow

Def. The excess-gap of each node $v \in V$ is $gap(v) \coloneqq f\left(\delta^{in}(v)\right) - b(v)$

Def. A node $v \in V$ is b-excessive (resp, b-deficient, b-balanced) if $gap(v) \; >$ (resp., $<$, $=$) 0.

Def. The feasibility gap of f is the total excess-gaps of all b-excessive nodes.

Fact. total excess-gaps of all nodes = 0, hence
total excess-gaps of all b-excessive nodes
= |total excess-gaps of all b-deficient nodes|

Gap-reducing augmenting path

s: b-excessive

t: b-deficient

P: an $s-t$ path in D_f

Fact. For any $0 < \varepsilon \leq |gap(s)| \wedge |gap(t)| \wedge \Delta_f(P)$,

- $f + \varepsilon\chi_P$ has smaller feasibility gap than f by ε
- no excessive node becomes deficient, and vice versa.

Recap. If f is extreme and P is a shortest $s-t$ path, then $f + \varepsilon\chi_P$ is also extreme.

Generic gap-reducing method

Goldberg-Tarjan [1988]:

$f \leftarrow 0$;

while f is not a b-TS

pick a b-excessive node s;

pick a b-deficient node t;

find a shortest s-t path P in A_f;

pick $0 < \varepsilon \leq |gap(s)| \wedge |gap(t)| \wedge \Delta_f(P)$;

$f \leftarrow f + \varepsilon \chi_P$;

return f

f is an extreme pseudoflow, and its feasibility gap is strictly decreasing.

Excess-gap scaling

Intuition: Choose a shortest but wide gap-reducing augmenting path
Analogy: **Edmonds-Karp capacity-scaling** algorithm for max-flow

b is integral

Excess-gap scaling factor Δ: a nonnegative integer power of 2

$$S(\Delta) \coloneqq \{v \in V : gap(v) \geq \Delta\}$$
$$T(\Delta) \coloneqq \{v \in V : gap(v) \leq -\Delta\}$$

Initial $\Delta \coloneqq 2^{\lfloor \log \quad \rfloor}$ where

$$B \coloneqq \min\left\{ \max_{v:b(v)>0} |b(v)|, \max_{v:b(v)<0} |b(v)| \right\}$$

so that either $S(2\Delta)$ or $T(2\Delta)$ is ∅

Δ-Scaling Phase

Invariant: At the beginning of the phase, either $S(2\Delta)$ or $T(2\Delta)$ is $\emptyset$

while $S(\Delta) \neq \emptyset$ and $T(\Delta) \neq \emptyset$
    pick $s \in S(\Delta)$ and $t \in T(\Delta)$;
    $P \leftarrow$ a shortest s-t path P in A_f;
    $f \leftarrow f + \Delta\chi_P$;

Divisibility: The pseudoflow carried on each edge and the excess-gap of every node are integer multiples of Δ

- $\Delta \leq |gap(s)| \wedge |gap(t)| \wedge \Delta_f(P)$;

At the end of 1-scaling phase, both $S(1)$ and $T(1)$ are $\emptyset$, and f is a b-TS

Distance-based potentials

- Initial potential for $f = 0$: Bellman-Ford or Dijkstra
- Each subsequent augmentation uses the shortest-path length as the potential

Running Time

Claim. There are at most $2n$ augmentations per scaling phase.

Pf. Δ-phase:

- In the beginning, the feasibility gap $< n(2\Delta) = 2n\Delta$.
- Each augmentation decreases the feasibility gap by Δ. ▪

Totally, $O(n \log B)$ augmentations, each using Dijkstra's algorithm.

Overall running time: $O((n \log B)((m + n \log n))$

Enhanced excess-gap scaling algorithm

[Orlin 1988,1993]

- b is fractional
- strongly polynomial time: $O((n \log n)((m + n \log n))$
- after each phase, Δ is decreased by a factor either 2 or $> 8n$
- an excessive node may become deficient, and vice versa
- amortized analysis of running time.

Overview of algorithm design

$f \leftarrow 0$, $\Delta := \max_{v \in V} |b(v)|$;

while f is not a b-TS

// Δ-phase

abundant-path augmentations;

shortest-path augmentations;

// update Δ

if $\max_{v \in V} |gap(v)| = 0$ then return f;

if $\max_{v \in V} |gap(v)| > \Delta/(8n)$ then $\Delta := \Delta/2$;

else $\Delta := \max_{v \in V} |gap(v)|$.

return f

Fact. At the beginning of Δ-phase, $\max_{v \in V} |gap(v)| > \Delta/(4n)$

Max-gap invariant in Δ-phase

initial max-gap $\leq 2(1-1/n)\Delta$
final max-gap $\leq (1-1/n)\Delta$

Def. A node v is Δ-large if $|gap(v)| > (1-1/n)\Delta$

Initial max-gap invariant is satisfied

- in the first phase: max-gap $= \Delta$.
- in each subsequent phase due to the final max-gap invariant in the previous phase.

Final max-gap invariant is ensured (easily) by the SP-augmentations.

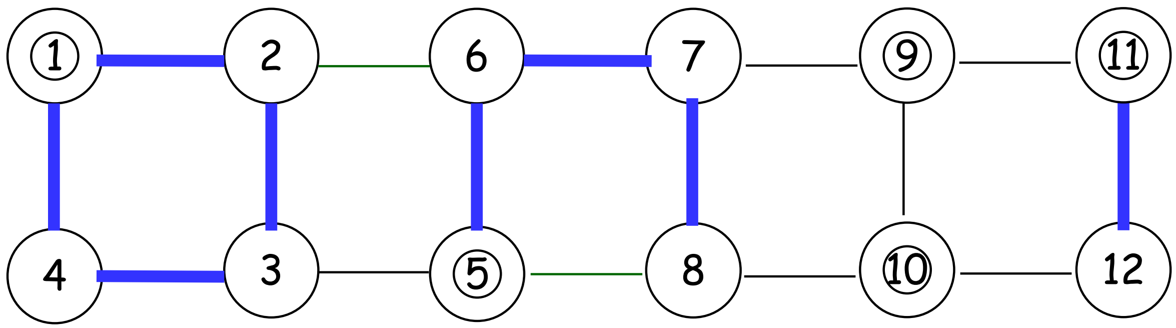
Abundant edges and components

Initially, all arcs are non-abundant

At the beginning of the Δ-phase, all nonabundant arcs e with

$$f(e) \geq 8n\Delta$$

are marked abundant.

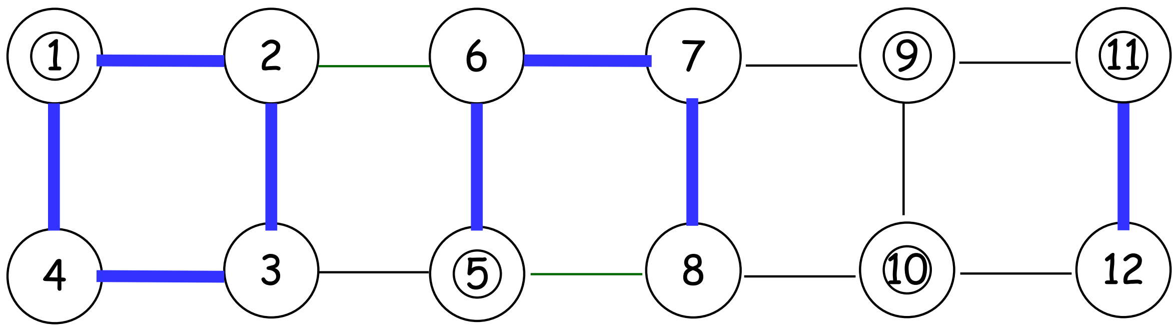


abundant subgraph : (V, abundant arcs).
abundant components: weak components in the abundant subgraph
roots: the minimum index node in an abundant component

Divisibility invariant in Δ-phase

flow on each nonabundant arc is an integer multiple of Δ

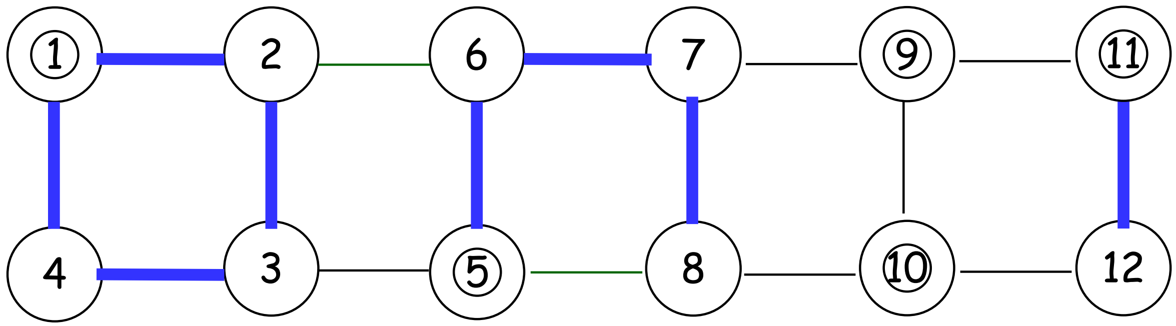
- satisfied trivially by $f = 0$
- ensured (easily) by the SP-augmentations and resetting of Δ



Abundance invariant in Δ-phase

for all abundant arcs e,

- initial full-abundance: $f(e) \geq 8n\Delta$
- subsequent half-abundance: $f(e) \geq 4n\Delta$



Initial full abundance is satisfied

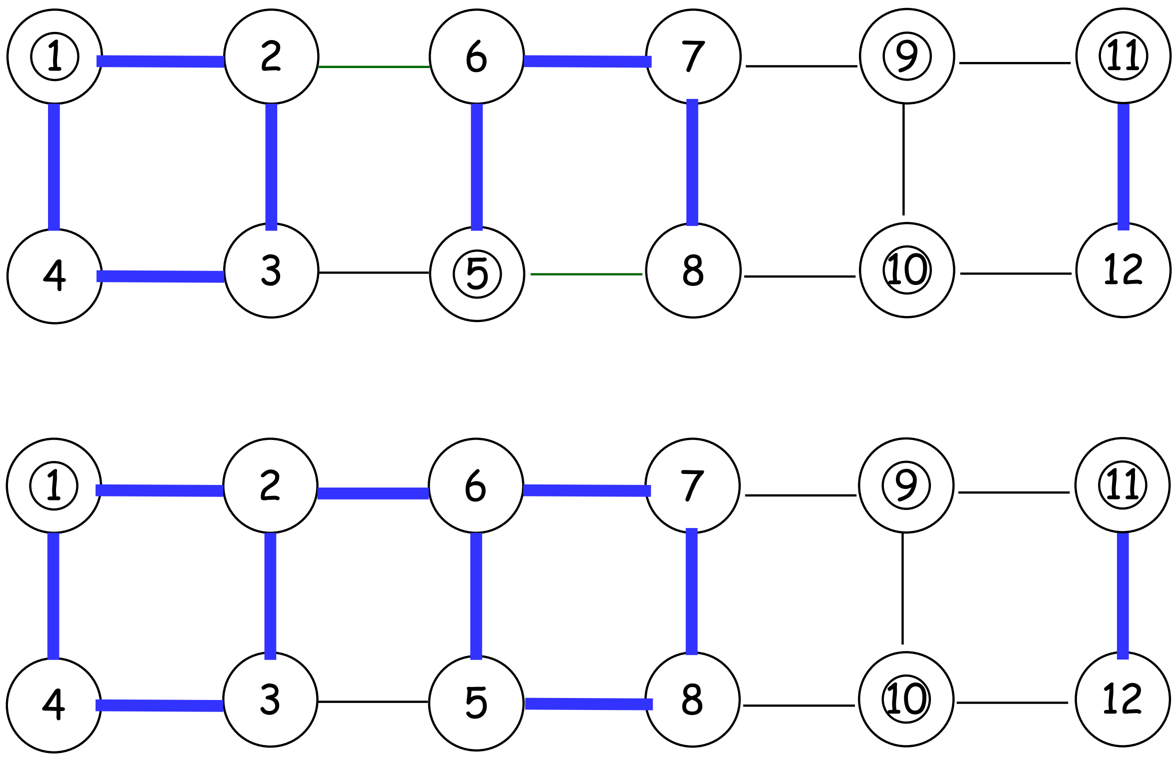
- in the first phase: trivial
- in each subsequent phase: new abundant arcs and final half-abundance in the previous phase.

Half-abundance is ensured jointly by AP & SP augmentations.

Merging of components

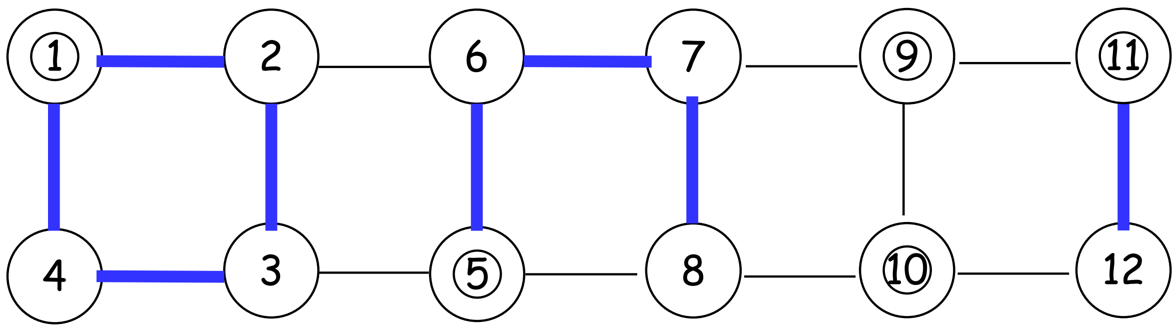
- Occurs after additional edges are marked abundant.
- Some roots before merging become non-roots after merging.

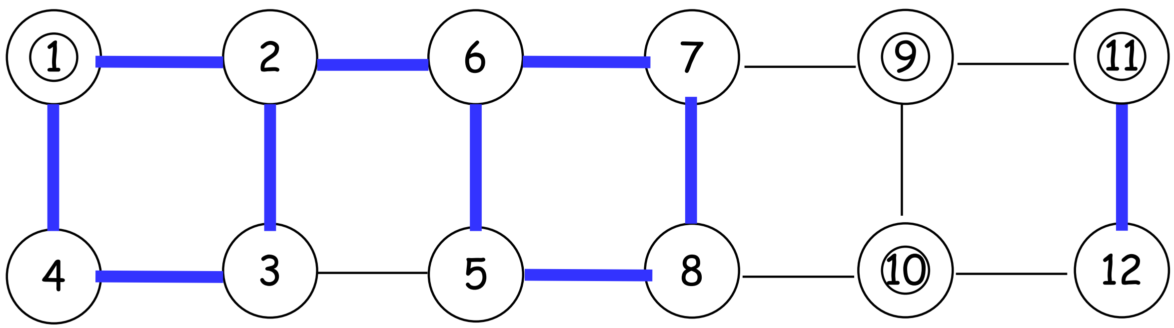
$n' :=$ number of new non-roots $=$ drop on the # of components



Excursion: gap shift

- Inner shift:
  - make all non-roots balanced
  - abundant-path augmentations
- Outer shift:
  - eliminate Δ-large nodes
  - shortest-path augmentations





Inner gap shift

n' inner pairs (s, t): new non-root t and its root s

```
for each inner pair (s, t)
    gap(t) ← 0;
    gap(s) ← gap(s) + gap(t);
```

The shift for (s, t) is

- constructive if $|gap(s) + gap(t)| \leq gap(s)$
- destructive if $|gap(s) + gap(t)| > gap(s)$

Eligible root-pairs

Def. A root-pair (s, t) is Δ-eligible if

- either $gap(s) > (1 - 1/n)\Delta$ and $gap(t) < -\Delta/n$
- or $gap(s) > \Delta/n$ and $gap(t) < -(1 - 1/n)\Delta$

Fact: There is an Δ-eligible root-pair $\Leftrightarrow$ there is a Δ-large node

Outer gap shift

$L \leftarrow \emptyset$;
while there is a Δ-large node
    $(s, t) \leftarrow$ a Δ-eligible root-pair;
    append (s, t) to L;
    $gap(s) \leftarrow gap(s) - \Delta$;
    $gap(t) \leftarrow gap(t) + \Delta$,

- s may become deficient but $|gap(s)| < (1 - 1/n)\Delta$
- t may become excessive but $|gap(t)| < (1 - 1/n)\Delta$
- no Δ-large nodes in the end

Counting pairs in L

$n'' \coloneqq$ # of remaining roots which are Δ–large before inner shift

Lemma. $|L| \leq n'' + 2n'$

Charging scheme:

- Each pair charges 1 coin to a Δ-large root
- For each root v receiving at least one charge:
  - first collects 2 coins from each new non-root in its component;
  - if not covering the charge, then pays 1 coin by its own

Claim. charges on $v \leq$ coins collected by v; and if v pays 1 coin itself, then it is Δ–large before inner shift

total # of coins collected $\leq n'' + 2n'$

Counting pairs in L

Claim. charges on $v \leq$ coins collected by v; and if v pays 1 coin itself, then it is Δ–large before inner shift

Pf. Let $k \coloneqq \#$ of new non-root nodes in its component. The inner shift for each of them increases $|gap(v)|$ by $\leq 2(1 - 1/n)\Delta$.

Case 1 : v is not Δ–large before inner shift. After inner shift,

$$|gap(v)| \leq (2k + 1)(1 - 1/n)\Delta.$$

Hence, v can be Δ–large for $\leq 2k$ times.

Case 2 : Otherwise. After inner shift,

$$|gap(v)| \leq 2(k + 1)(1 - 1/n)\Delta.$$

Hence, v can be Δ–large for $\leq 2k + 1$ times.

Total gap shift

total inner gap shift $\leq 2n'(1 - 1/n)\Delta \leq 2n'\Delta$

total outer gap shit $\leq (n'' + 2n')\Delta$

total gap shift $\leq 2n'\Delta + (n'' + 2n')\Delta = (n'' + 4n')\Delta \leq 4n\Delta$

Abundant-path augmentations

for each inner pair (s, t)
  $P \leftarrow$ an abundant s-t path;
  augment f along P by $|gap(t)|$ units;

- Half-abundance & divisibility are maintained
- No change to A_f, hence f is extreme throughout
- Running time: $O(n'm)$

Shortest-path augmentations

while there is a Δ-large node
    $(s, t) \leftarrow$ a Δ-eligible root-pair;
    $P \leftarrow$ a shortest s-t path in A_f;
    $f \leftarrow f + \Delta\chi_P$;

- Half-abundance & divisibility are maintained
- Each P has bottleneck capacity $\geq \Delta$
- f is extreme throughout

Divisibility after resetting Δ

Δ': the scaling factor in the next phase

<u>Case 1</u>: $\Delta' = \Delta/2$. Trivial.

<u>Case 2</u>: $\Delta' < \Delta/2$. All arcs e with $f(e) > 0$ become abundant:

$$f(e) \geq \Delta \geq 8n\Delta'.$$

Hence all non-abundant arcs carry 0 flow.

Running time

Thm. There are $O(n \log n)$ scaling phases and $O(n \log n)$ shortest-path augmentations.

Running time:

- merging and abundant-path augmentations: $O(nm)$
- shortest-path augmentations: $O((n \log n)((m + n \log n))$
- resetting Δ: $O(n^2 \log n)$

Overall: $O((n \log n)((m + n \log n))$

Component merging

S: a component with $|gap(S)| \geq \Delta/(4n)$ at the beginng of the Δ-phase.

Lemma. Within $k \coloneqq \lceil \log(36n^2m) \rceil = O(\log n)$ additional scaling phases, S will be merged into a larger component.

Claim. $|b(S)| \geq \Delta/(4n)$ at the beginng of the Δ-phase.

Pf. If $f\left(\delta^{in}(S)\right) = 0$ then $|b(S)| = |gap(S)| \geq \Delta/(4n)$.
Otherwise,

$$|b(S)| \geq \left|f\left(\delta^{in}(S)\right)\right| - |gap(S)| \geq \Delta - 2(1 - 1/n)\Delta > \Delta/(4n).$$

By contradiction. Otherwise, within k scaling phases, the scaling factor will be $\Delta' \leq \Delta/(36n^2m)$, or $\Delta/(4n) \geq 9nm\Delta'$.

Component merging

Claim. At the beginning of the Δ'-phase, $\delta(S)$ contains an abundant edge.

Pf.

$$\left|f\left(\delta^{in}(S)\right)\right| \geq |b(S)| - |gap(S)| \geq \Delta/(4n) - 2(1 - 1/n)\Delta'$$
$$> 9nm\Delta' - 2\Delta' \geq 8nm\Delta'.$$

Hence $\delta(S)$ contains an e with $f(e) > 8n\Delta'$.

After AP-augmentations, S gets merged into a larger component, a contradiction.

Counting scaling phases

Thm. There are $O(n \log n)$ scaling phases.

Pf.
S: a component with largest $|gap(S)|$ at the beginng of the Δ-phase.

$$|gap(S)| \geq \Delta/(4n)$$

S will be merged into a larger component within $O(\log n)$ scaling phases.

Counting shortest-path augmentations

Thm. There are $O(n \log n)$ shortest-path augmentations.

Charging + coin collection scheme

- Each new non-root pays ≤ 2 coins $\Rightarrow \leq 2n$ coins from new non-roots.
- Each root v pays $O(\log n)$ coins before its component S gets merged.

Pf. Suppose v pays > 1 coins before S gets merged. At the beginning of the Δ-phase where v pays its second coin,

$$|gap(S)| = |gap(v)| > (1 - 1/n)\Delta > \Delta(4\mathrm{n}).$$

$\leq 2n$ different components $\Rightarrow \leq O(n \log n)$ coins from roots

3. Min-Cost Circulation

Generic Circuit Canceling Method

Idea:

- Maintain a feasible circulation (capacity & excess constraints).
- Progress toward optimality (no negative circuit)

Klein [1967]:

$f \leftarrow$ an initial circulation;
while there is a negative circuit in A_f
  find a negative circuit C in A_f;
  $f \leftarrow f \oplus C$;
return f

Min-Mean Circuit Canceling

Goldberg-Tarjan [1988]

- Each iteration finds a min-mean circuit in time $O(mn)$.
- Conceptually simple yet strongly polynomial

Thm. Total number of iterations is $O(m^2 n \log n)$.

- $f_0, f_1, f_2, \dots$: sequence of circulations
- $A_i \coloneqq A_{f_i}$

For all i except the last one,

- $C_i \coloneqq$ min-mean circuit in A_i
- $\mu_i \coloneqq -\ell(C_i)/|C_i|$: "optimality gap"
- C_i^*: the set of bottleneck edges in C_i

$$A_{i+1} = [(A_i \setminus C_i^*) \cup C_i^{-1}]$$

Overview of the analysis

$$l := mn\lceil \ln(2n-1) \rceil$$

C_i has an edge a s.t. neither a nor a^{-1} appears in any C_j with $j \geq i + l$.
$\Rightarrow$ total number of iterations $O(ml)$

Decreasing sequence: $\mu_0 \geq \mu_1 \geq \mu_2 \geq \cdots$

- after m iterations: $\mu_{i+m} \leq (1 - 1/n)\mu_i$
- after mn iterations: $\mu_{i+mn} \leq (1 - 1/n)^n \mu_i < \mu_i / e$
- after l iterations: $\mu_{i+l} < \mu_i / e^{\ln(2n-1)} = \mu_i/(2n-1)$.

Recap: Adjusted edge lengths (prices)

- Raising the length of each edge in A_i by μ_i makes all circuits in A_i non-negative.

- Node price function p_i: $p_i(v) - p_i(u) \leq \ell(u,v) + \mu_i, \forall (u,v) \in A_i$

- ℓ_i: p_i-adjusted edge length function on A.

$$\ell_i(u,v) \coloneqq \ell(u,v) + p_i(u) - p_i(v), \forall (u,v) \in A$$

- for any circuit C in A, $\ell_i(C) = \ell(C)$
- $\ell_i(a) \geq -\mu_i$ for each $a \in A_i$; $\ell_i(a) = -\mu_i$ for each $a \in C_i$
  - each $a \in A$ with $\ell_i(a) < -\mu_i$ is not in A_i, i.e., $f_i(a) = c(a)$
  - $\ell_i(a) \leq \mu_i$ for each $a \in A_i^{-1}$; $\ell_i(a) = \mu_i$ for each $a \in C_i^{-1}$

Frozen effect

$$C_i' \coloneqq \{a \in C_i : \ell_{i+l}(a) \leq -\mu_i\}$$

- $C_i' \neq \emptyset$ since $\ell_{i+l}(C_i)/|C_i| = \ell(C_i)/|C_i| = -\mu_i$
- Each $a \in C_i'$ is extraordinarily cheap: $\ell_{i+l}(a) \leq -\mu_i < -(2n-1)\mu_{i+l}$

Thm. For each $a \in C_i'$, $f_j(a) = c(a)$ (i.e., $a \notin A_j$) for all $j \geq i + l$.

Coro. For each $a \in C_i'$, and $a^{-1} \notin C_j$ for all $j \geq i + l$ except the last one.

Pf.

$$\begin{aligned} a^{-1} \in C_j &\Rightarrow f_{j+1}(a^{-1}) > f_j(a^{-1}) \\ &\Rightarrow f_{j+1}(a) < f_j(a) = c(a) \end{aligned}$$

Frozen effect

Thm. For each $a \in C_i'$, $f_j(a) = c(a)$ (i.e., $a \notin A_j$) for all $j \geq i + l$.

Pf. $\ell_{i+l}(a) < -(2n-1)\mu_{i+l} < -\mu_{i+l} \Rightarrow a \notin A_{i+l} \Rightarrow f_{i+l}(a) = c(a)$.

Assume by contradiction $f_j(a) < c(a) = f_{i+l}(a)$ for some $j > i + l$. Then $A^+(f_{i+l} - f_j)$ contains a circuit C s.t. $a \in C \subseteq A_j \cap A_{i+l}^{-1}$.

- $a \in C \subseteq A_{i+l}^{-1} \Rightarrow$
$$\ell(C) = \ell_{i+l}(C) = \ell_{i+l}(a) + \ell_{i+l}(C \setminus \{a\})$$
$$< -(2n-1)\mu_{i+l} + (|C|-1)\mu_{i+l} = -(2n-|C|)\mu_{i+l} \leq -|C|\mu_{i+l}$$

- j is not the last one, and $C \subseteq A_j \Rightarrow \ell(C) = \ell_j(C) \geq -|C|\mu_j \geq -|C|\mu_{i+l}$

a contradiction!

Drop after single-iteration

Lemma. $\mu_{i+1} \leq \mu_i$

Pf.

$$C_{i+1} \subseteq A_{i+1} \subseteq A_i \cup C_i^{-1}$$
$$\Rightarrow \ell_i(a) \geq -\mu_i \text{ for each } a \in C_{i+1}$$
$$\Rightarrow \ell_i(C_{i+1}) \geq -|C_{i+1}|\mu_i$$
$$\Rightarrow \mu_{i+1} = -\ell(C_{i+1})/|C_{i+1}| = -\ell_i(C_{i+1})/|C_{i+1}| \leq \mu_i.$$

Geometric drop after m iterations

Lemma. $\mu_{i+m} \le (1 - 1/n)\mu_i$.

Pf.

$$C_i \subseteq A^- := \{a \in A : \ell_i(a) < 0\}$$
$$k := \max\{j \ge i \ : C_j \subseteq A^-\}$$

Claim. $A_j \cap A^-$ is strictly decreasing for $i \le j \le k+1$.

Pf. For each $i \le j \le k$, $\emptyset \ne C_j^* \subseteq C_j \subseteq A_j \cap A^-$, hence

$$A_{j+1} \cap A^- = [(A_j \setminus C_j^*) \cup C_j^{-1}] \cap A^- = (A_j \cap A^-) \setminus C_j^* \subset A_j \cap A^-.$$

Claim. $k < i + m$.

Pf. Otherwise,

$$|A^-| \ge |A_i \cap A^-| \ge |A_k \cap A^-| + k - i \ge |C_k| + m > m.$$

Geometric drop after m iterations

Claim. $\mu_{k+1} \leq (1 - 1/n)\mu_i$.

Pf.

$\ell_i(C_{k+1}) \geq \ell_i(C_{k+1} \cap A^-)$ // keeping only negative edges

$\geq -|C_{k+1} \cap A^-|\mu_i$ // $C_{k+1} \cap A^- \subseteq A_{k+1} \cap A^- \subset A_i \cap A^- \subseteq A_i$

$\geq -(|C_{k+1}| - 1)\mu_i$ // $C_{k+1} \nsubseteq A^-$ by the choice of k

Thus,

$$\mu_{k+1} = -\frac{\ell(C_{k+1})}{|C_{k+1}|} = -\frac{\ell_i(C_{k+1})}{|C_{k+1}|} \leq \left(1 - \frac{1}{|C_{k+1}|}\right)\mu_i \leq \left(1 - \frac{1}{n}\right)\mu_i.$$

Finally, since $i + m \geq k + 1$, $\mu_{i+m} \leq \mu_{k+1}$

Summary

- Shortest-path augmenting
  - Maintain extremeness (i.e., no negative circuit), and move toward excess feasibility
- Circuit canceling
  - Maintain excess feasibility, and move toward optimality (i.e., no negative circuit)
- Advanced variants
  - generalized flow: linear gain/loss
  - convex cost
  - submodular transshipment (flow): the excesses satisfy submodular constraints