CMSC 451 - Algorithm Design
Lecture 18 - Approximation: VC, TSP, + Bin Packing

Coping with NP-Completeness:

What to do if your problem is NP-hard?

- Brute-force search: Only for small inputs
- Heuristics: (e.g. greedy) Fast but no guarantees (unless you prove them)
- General search: Often the best choice
  - Branch + bound
  - Metropolis-Hastings
  - Simulated annealing
  - Genetic / Evolutionary
  - Local search
  - ...
- Approximations:
  - Guaranteed performance
  - Specialized for particular problem

How close to optimal?

Approximation ratio: Ratio of heuristic value to optimal = ρ (Greek rho)

Common approx. results: (Best to worst)

- Can be approximated arbitrarily well
  - User gives an allowed error bound $\varepsilon > 0$ and approx. ratio is $\leq 1+\varepsilon$
  - Running time poly. in n but grows with ε (e.g. $O((1/\varepsilon)^2 \cdot n^3) \rightarrow \infty$ as $\varepsilon \rightarrow 0$
  - Called Polynomial Time Approximation Scheme (PTAS - pronounced "P-tas")
  - Examples: Subset sum, Euclidean TSP
- Constant approx. ratio
  - Examples:
    - k-center (Gonzalez): $\rho = 2$
    - Vertex cover: $\rho = 2$
    - TSP: $\rho = 2$ (or better $\rho = 1.5$)
    - Bin packing: $\rho = 2$
- Logarithmic:
  - Example:
    - Greedy set cover: $\rho = \ln n$
    - Essentially best possible (unless P = NP)

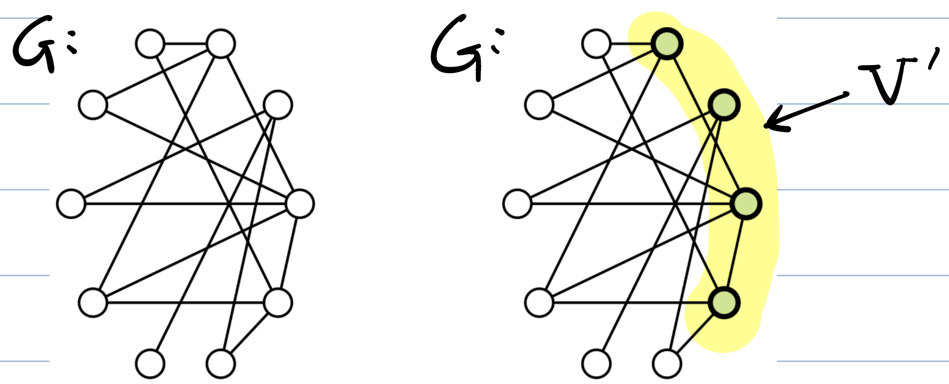
- Inapproximable: No poly. time algorithm achieves an approx. unless P = NP.

Vertex Cover Approximation:

- Present a simple approx. alg with $\rho = 2$

Vertex Cover (Optimization)

Given a graph $G=(V,E)$ compute the smallest subset $V' \subseteq V$ s.t. every edge is incident to some vertex in V'.



First Attempt: Greedy Heuristic

- Repeatedly select vertex u of highest degree + add to V'
- Remove u + all its incident edges
- Update vertex degrees

- Equivalent to greedy set cover on $(X, \mathcal{S})$

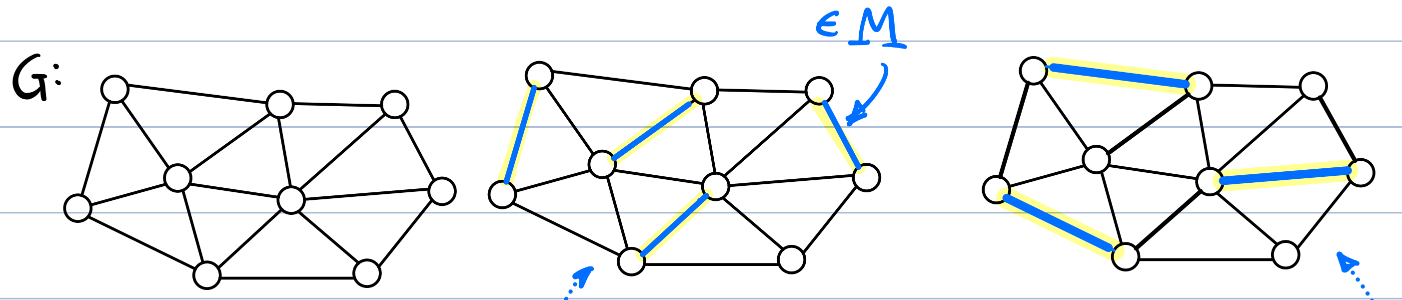
X = edge set = E

$\mathcal{S}$ = $\forall u \in V$, $S_u = \{ e \mid e \text{ incident to } u \}$

Approx. ratio = $\ln m$ $\quad m = |E|$

Second attempt: Based on maximal matching

- Recall: A matching in G is a subset of edges $M \subseteq E$ s.t. each vertex incident to at most one edge of M



- Maximum matching: largest size possible
- Maximal matching: cannot add more edges

- Maximum matchings are nontrivial to compute (Edmond's blossom algorithm)
- Maximal matchings are very easy

Key observation: M = any maximal matching
V^* = optimal vertex cover

Claim:

(i) $|V^*| \geq |M|$

(ii) $|V^*| \leq 2 \cdot |M|$

Proof:

(i) For any $(u,v) \in M$, either u or v must be in V^* (BTW, This holds for any matching and any vertex cover.)

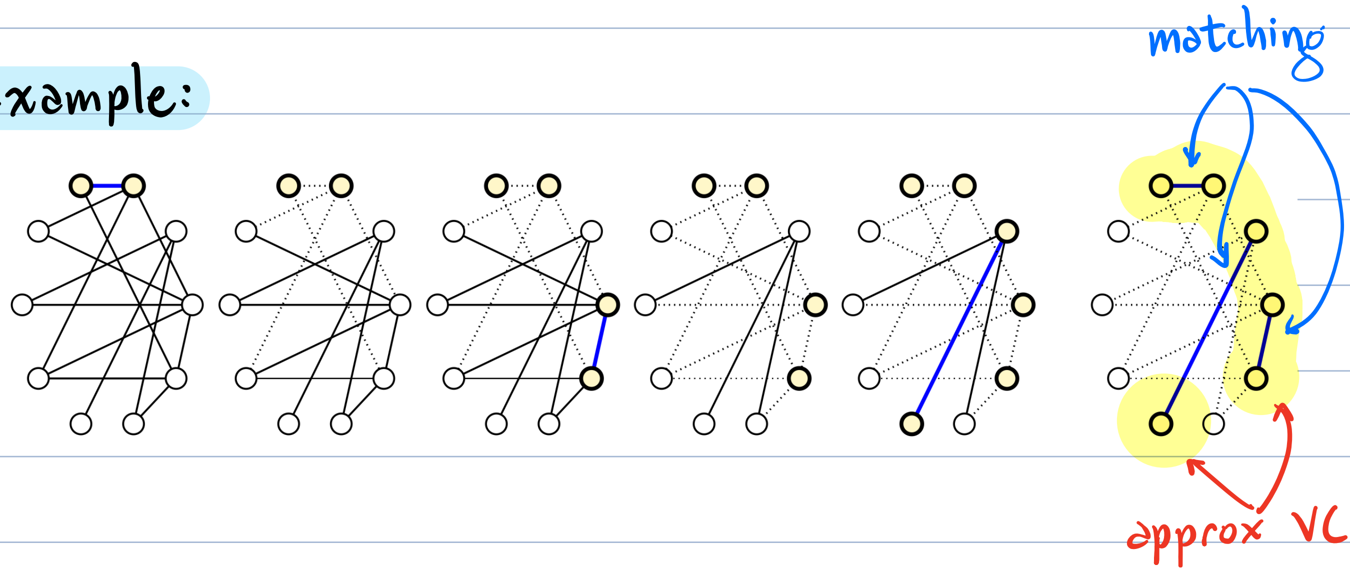
(ii) Since M is maximal, for every $(u,v) \in E$, either u or v is incident to an edge of M (otherwise, we could add (u,v) to M).

$\Rightarrow$ Taking all endpts of M yields a vertex cover, of size $2 \cdot |M|$. V^* can't be larger. $\square$

Approach: Compute any maximal matching M + for each $(u,v) \in M$, add both u+v to V'.

```
approx-VC(G)
    V' ← ∅
    while (G has an edge, (u,v))        // take an edge
        V' ← V' ∪ {u,v}                 // add it end pts
        delete u + v and incid. edges
```

Example:



Claim: approx-VC achieves approx. ratio of 2

Proof: V' is a vertex cover.

$|V'| = 2 \cdot |M|$, where M is maximal match.

By Lemma (i) $|M| \le |V^*|$

$\Rightarrow |V'| = 2|M| \le 2|V^*|$

Reductions + Approximations:

- Since VC has a 2-factor approximation don't all NP problems (via reduction) have a 2-factor approximation?

- No! Reductions don't necessarily preserve approx. ratios.

- Case study: $IS \leq_p VC$ (we showed:)

G has vertex cover of size k

iff

G has indep. set of size n-k

- Let k^* = opt. VC size.

$\Rightarrow$ We can compute a vert. cover of size $\leq 2k^*$

$\Rightarrow$ can compute an indep. set of size $\geq n-2k^*$

- Approx. ratio for IS:

$$\frac{\text{opt}}{\text{approx}} \leq \frac{n-k^*}{n-2k^*} \longleftarrow \text{Could be huge!}$$

E.g. $n = 1{,}001$, $k^* = 500$, $n-k^* = 501$

$n - 2 \cdot k^* = 1$

$$\frac{n-k^*}{n-2\cdot k^*} = \frac{501}{1} = 501 \text{ !}$$

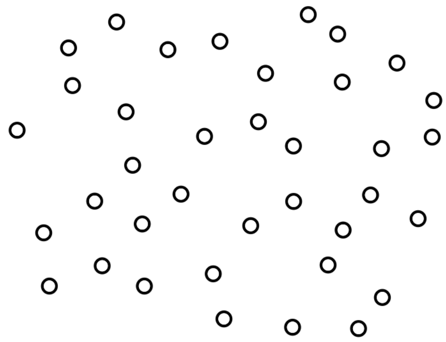
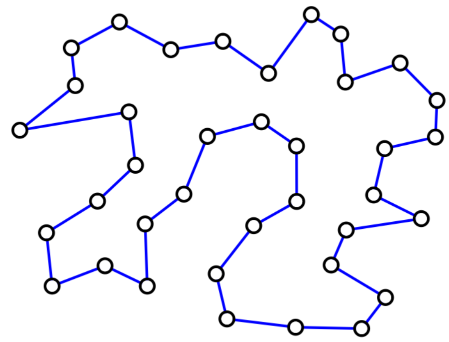
Traveling Salesman Problem (Metric case)

Recall: A (discrete) metric space is a finite point set P + distance function $d(u,v)$ $\forall\, u,v \in P$, where:

- $d(u,v) \geq 0$ + $d(u,v) = 0$ iff $u = v$
- $d(u,v) = d(v,u)$
- $d(u,w) \leq d(u,v) + d(v,w)$ (Triangle inequality)

- Most natural dist. functions are metrics:
  - Euclidean dist
  - Shortest path dist. in graphs
  - Edit dist. (Levenshtein dist.)
- Representable as:
  - $n \times n$ distance matrix (symmetric)
  - complete weighted graph

(Metric) Traveling Salesman Problem (TSP): Given a set P of n points in a metric space, compute a simple cycle that visits all vertices of minimum weight.

Example: P:  $TSP(P)$: 

Given a set T of edges, let $wt(T)$ = sum of distances.

Obs: Given any TSP tour, remove an edge. A spanning tree remains $\Rightarrow$

$$wt(MST(P)) \leq wt(TSP(P))$$

Idea: If we can compute a valid tour H of weight $\leq c \cdot wt(MST(P))$ then:

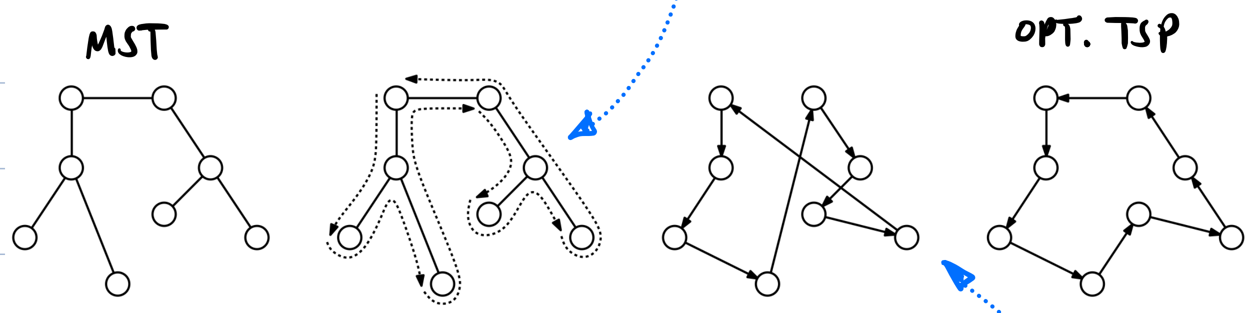
$$wt(H) \leq c \cdot wt(MST(P)) \leq c \cdot wt(TSP(P))$$

$\Rightarrow$ H is a factor-c approximation!

How to convert MST into a cycle?

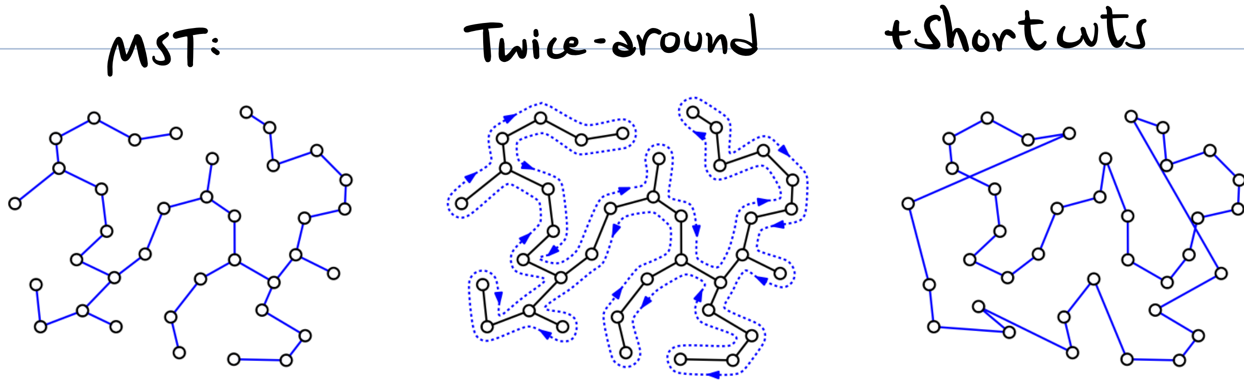
The Twice-Around Tour:

- Start at any vertex of MST
- Hold your hand against the MST "wall"
  walk around the MST



- This cycle hits each edge twice, but
  repeats points, not allowed.
- Shortcut the path by removing repeats.
- By triangle inequality - shortcutting
  can only reduce lengths.

Example:



Since graph is complete, can compute MST in $O(n^2 \log n)$ time. Twice-around + shortcuts add $O(n)$ time.

Lemma: The Twice-Around alg. has an approx. ratio of 2.

Proof: Let H = twice-around tour &
let H^* = opt TSP tour &
let T = MST edges

Earlier:

$$wt(T) \le wt(H^*)$$

Since "pure" twice around hits each edge twice & shortcutting only decreases weights:

$$wt(H) \le 2 \cdot wt(T)$$

$\Rightarrow$

$$wt(H) \le 2 \cdot wt(T) \le 2 \cdot wt(H^*) \Rightarrow \frac{wt(H)}{wt(H^*)} \le 2$$

$\square$

Christofides Algorithm:

- A $\frac{3}{2}$-factor approx to metric TSP
- By Nicos Christofides (1976)

Obs: Wastage in twice-around comes from visiting each MST edge twice.

- Can we avoid this repetition?

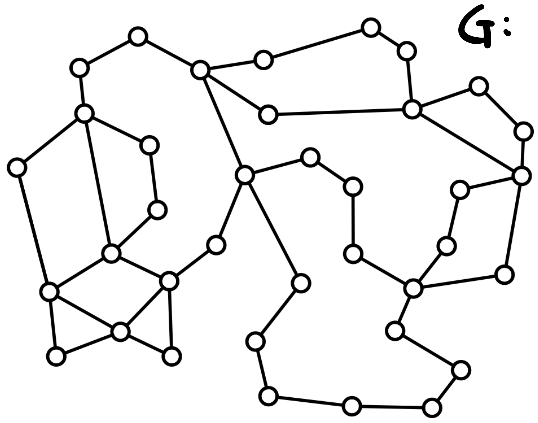
Eulerian Circuit: A cycle in a graph that visits every edge exactly once

- The following is well known:

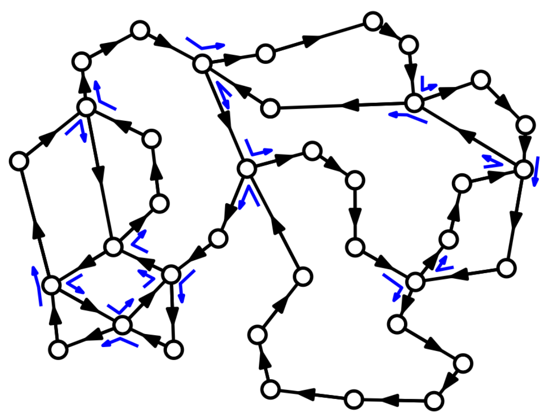
> **Lemma:** A graph $G = (V, E)$ has an Eulerian circuit iff every vertex has even degree. Can be computed in $O(n+m)$ time.

Example:

G:

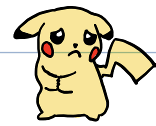


Eulerian Circuit



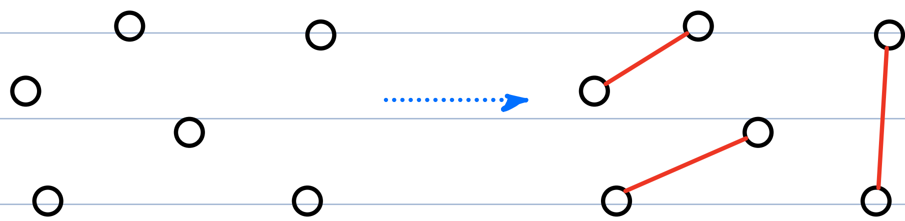
How to exploit this?

- MST definitely has vertices of odd degree (leaves have degree 1)



Christofides idea:

- Extract all vertices of MST of odd degree
- Fact: Num. of vertices of odd degree is even (Handshaking lemma)
- Compute a matching on these vertices



- Add there edges to MST
  $\Rightarrow$ now all vertices have even degree!
- Use Eulian circuit rather than twice around

Background:

Matching: Set of edges, each vertex incident to ≤ 1

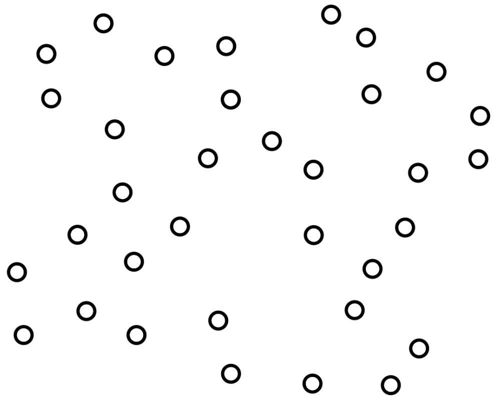
Perfect mething: Every vertex incident to 1 edge

Min-weight perfect matching (MWM(P))

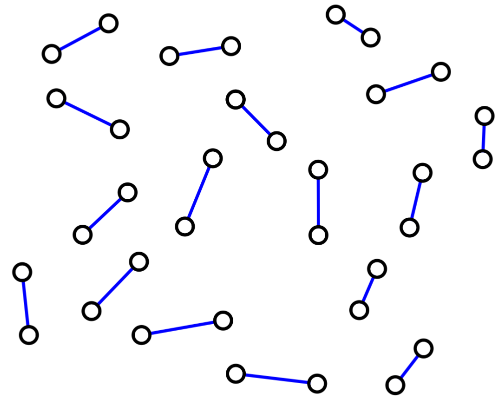
- perfect matching of min weight
- Computable in $\mathcal{O}(|V|^2 \cdot |E|)$

Assume $|P|$ even

Example: P:



$MWM(P)$



Christofides Algorithm: Given P in metric space.

(a) Compute edges T of $MST(P)$

(b) Let $P_1 \leftarrow$ pts of P with odd deg. in $MST(P)$

(c) Compute edges M of $MWM(P_1)$

(d) Form graph $G = (P, T \cup M)$ [G has even deg.]

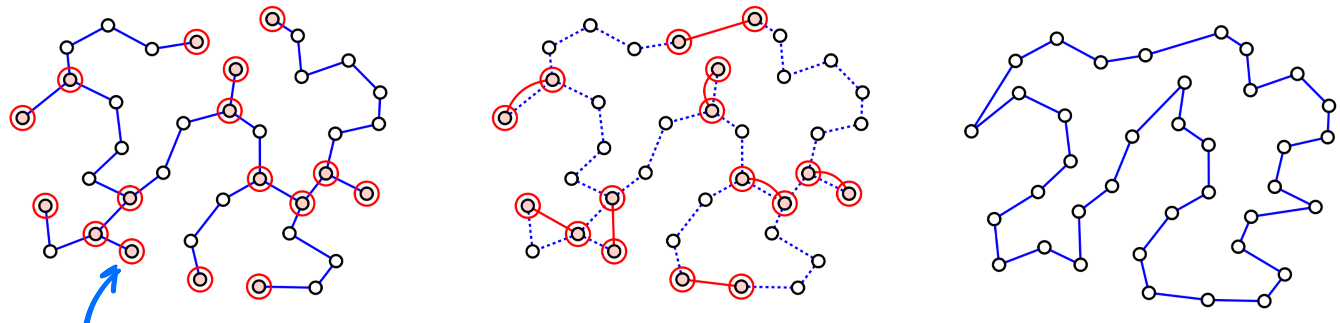
(e) $C' \leftarrow$ Eulerian circuit in G

$C \leftarrow$ apply short-cutting to C' [$wt(C) \leq wt(C')$]

(f) return C

Example:

MST(P) MWM(P_1) Eulerian circuit + shortcuts



P_1 = odd deg vertices

- Using an Eulerian circuit is more cost efficient, but how much weight does MWM(P_1) add?

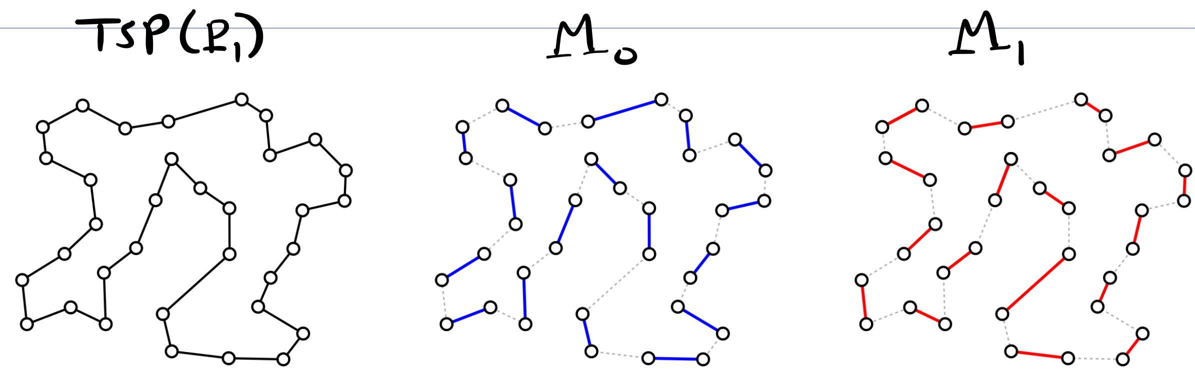
Lemma: Given point set P, let P_1 be any subset of even cardinality. Then:

$$wt(MWM(P_1)) \leq \frac{wt(TSP(P))}{2}$$

Proof: Removing pts. only decreases TSP:

$$wt(TSP(P_1)) \leq wt(TSP(P))$$

- Split $TSP(P_1)$ into alternating groups M_0 + M_1



$\Rightarrow$ $wt(TSP(P_1)) = wt(M_0) + wt(M_1)$

- Both M_0 + M_1 are perfect matchings (but not nec. minimum weight)

$\Rightarrow$ $wt(MWM(P_1)) \leq \min(wt(M_0), wt(M_1))$

- Min of two numbers $\leq$ average

$$\Rightarrow \min(wt(M_0), wt(M_1)) \leq \frac{wt(M_0) + wt(M_1)}{2}$$

- Combining:

$$\begin{aligned} wt(MWM(P_1)) &\leq \min(wt(M_0), wt(M_1)) \\ &\leq \frac{wt(M_0) + wt(M_1)}{2} \\ &= wt(TSP(P_1))/2 \leq \frac{wt(TSP(P))}{2} \quad \square \end{aligned}$$

Lemma: Christofides alg. has approx ratio $3/2$

Proof: Let: $H^* =$ opt TSP tour on P

$T = MST(P)$

$M = MWM(P_1)$

As shown earlier: $wt(T) \leq wt(H^*)$

By previous lemma:

$$wt(M) \leq \frac{1}{2} wt(H^*)$$

Let C' be Euler circuit in $T \cup M$

& $C = \text{shortcut}(C')$

We have:

$$\begin{aligned} wt(C) \leq wt(C') &= wt(T) + wt(M) \\ &\leq wt(H^*) + \frac{1}{2} wt(H^*) \\ &\leq \frac{3}{2} wt(H^*) \end{aligned}$$

$$\Rightarrow \frac{wt(C)}{wt(H^*)} \leq \frac{3}{2}$$

$\square$

Summary -

- Approximation ratios
- Vertex cover ($\rho = 2$)
- Metric TSP
  - Twice-around ($\rho = 2$)
  - Christofides ($\rho = 3/2$)