No External Regret Algorithm

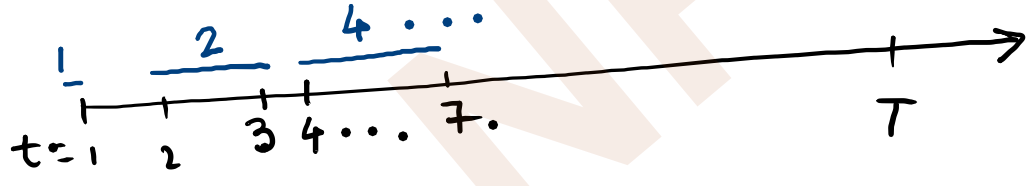
Lecture 7.1

Theorem: $|A| = n$. There exists an algorithm with external regret (time-averaged) $O\left(\sqrt{\frac{\log n}{T}}\right)$.

Multiplicate weight algorithm. Set $\varepsilon = \sqrt{\frac{\ln n}{T}}$. The problem with setting $\varepsilon = \sqrt{\frac{\ln n}{T}}$ is that we need to know the time horizon T in advance.

Remark: $|A| = n$. Then there exists an algorithm with (time-averaged) external regret $O\left(\sqrt{\frac{\ln n}{T}}\right)$. Moreover the algorithm does not need to know T a-priori.

Proof:



Group the iretations into epochs. ξ_i, $i = 0, \ldots, l$

$$\xi_i = \left[2^i, \min\{2^{i+1} - 1, T\} \right]$$

In the beginning of each epoch, we reset the weight vector of the MW algorithm to the all 1's vector.

In the i-th epoch, we use $\varepsilon = \varepsilon_i = \sqrt{\frac{\ln n}{2^i}}$. Let OPT_i be the pay off of any fixed action in the i-th epoch.

$$OPT - \sum_{i=1}^{T} r_i \leq \sum_{i=0}^{l} \left[OPT_i - \sum_{t \in \xi_i} r_i \right]$$

$$\leq \sum_{i=0}^{l} \left[\varepsilon_i \cdot 2^i + \frac{\ln n}{\varepsilon_i} \right]$$

$$= \sqrt{\ln n} \sum_{i=0}^{l} 2^{\frac{i}{2}+1}$$

$$\leq 2^{\frac{l}{2}+2} \cdot \sqrt{\ln n}$$

$$\leq 4\sqrt{T \ln n}$$

$$\frac{1}{T}\left(OPT - \sum_{i=1}^{T} \gamma_i\right) \leq 4\sqrt{\frac{\ln n}{T}}$$

Combining Expert Advice: The problem of designing a no-regret algorithm is also called "combining expert

"advice" - the mixed strategy of every iteration is like "combining experts' advices" and the goal is to perform as good as an expert asymptotically.

Combining exploration and exploitation.

Connection between no-external-regret dynamic and equilibrium concepts:

Let $\Gamma = \langle N, (S_i)_{i \in N}, (u_i)_{i \in N} \rangle$ be any normal form game.

Each player runs a no-external regret algorithm.

$A_i = S_i$

$p_t \in \Delta(S_i)$

$$\pi_i(s_i) = u_i(s_i, p_{-i}) = \sum_{s_{-i} \in S_{-i}} \left[u_i(s_i, s_{-i}) \prod_{\substack{j \in [n] \\ j \neq i}} p_j(s_j) \right]$$

Theorem: Let $\varepsilon > 0$, players run their no-external-regret algorithm for T iterations such that their time-averaged external regret is at most ε. Define $\sigma_i = \prod_{i \in N} p_i^t$, $\sigma = \frac{1}{T} \sum_{i=1}^{T} \sigma_i$. Then σ is an ε-CCE of T.

That is $\mathbb{E}_{s \sim \sigma} [u_i(s)] \geq \mathbb{E}_{s \sim \sigma} [u_i(s_i', s_{-i})] - \varepsilon \quad \forall i \in N, \forall s_i' \in S_i$

Lecture 7.2

Theorem: Let $\varepsilon > 0$. Let T be such that the time-averaged external regret of every player is at most ε.

Define $\sigma_t = \prod_{i \in N} p_i^t$ and $\sigma = \frac{1}{T} \sum_{t=1}^{T} \sigma_t$. Then σ is an ε-CCE.

Proof:

$$\sigma = \frac{1}{T} \sum_{t=1}^{T} \sigma_t$$

$$\mathbb{E}_{s \sim \sigma}[u_i(s)] = \frac{1}{T} \sum_{t=1}^{T} \mathbb{E}_{s \sim \sigma_t}[u_i(s)]$$

For every $i \in N$, every $s_i' \in S_i$

$$\frac{1}{T}\left[\sum_{t=1}^{T} \mathop{\mathbb{E}}_{(s_i, s_{-i}) \sim \sigma_t} \left[u_i(s_i', s_{-i}) \right] - \sum_{t=1}^{T} \mathop{\mathbb{E}}_{s \sim \sigma_t} \left[u_i(s) \right] \right] \leq \varepsilon$$

$$\mathop{\mathbb{E}}_{(s_i, s_{-i}) \sim \sigma} \left[u_i(s_i', s_{-i}) \right] - \mathop{\mathbb{E}}_{s \sim \sigma} \left[u_i(s) \right] \leq \varepsilon$$

Hence, σ is an ε-CCE.

Q. Does there exist similar natural learning dynamic which enables player to converge to a CE?

Swap-Regret:

"Modification rule" $(\pi_1, \ldots, \pi_t, p_1, \ldots, p_t, a_1, \ldots, a_t)$ history. takes the history of the play as input and outputs a sequence of action for t iterations.

External regret tries to perform as good as any fixed action; i.e. the modification rules $\{F_a = (a, \ldots a) \mid a \in A\}$ (T times)

Swap regret tries to perform as good as any modification rule $\{F_{a,b} : a, b \in A, a \neq b\}$

↑ whenever the player played a, it outputs b and vice versa.

No Swap-Regret: A learning algorithm is said to have no swap regret if the time-averaged swap regret

$$\frac{1}{T}\left(\max_{\delta \in F^{swap}} \sum_{t=1}^{T} \sum_{a \in A} p_t(a) \cdot \pi_t(\delta(a)) - \sum_{t=1}^{T} \sum_{a \in A} p_t(a)\, \pi_t(a)\right)$$

goes to 0 as T goes to ∞.

- The connection of no swap-regret algorithm and CE:

Theorem: Let $\varepsilon > 0$. Each player runs a no-swap-regret algorithm.

$$A_i = S_i, \quad \pi_i(s_i) = u_i(s_i, p_{-i}) = \sum_{(s_j)_{j \neq i} \in S_{-i}} u_i(s_i, s_{-i}) \cdot p_j(s_j)$$

Let T be such that the time-averaged swap-regret of every player is at most ε. Define $\sigma_t = \prod_{t=1}^{T} p_t$ and $\sigma = \frac{1}{T} \sum_{t=1}^{T} \sigma_t$. Then σ is an ε-CE of T.

Proof: By definition of σ,

$$\mathop{\mathbb{E}}_{s\sim\sigma}[u_i(s)] = \frac{1}{T}\sum_{t=1}^{T} \mathop{\mathbb{E}}_{s\sim\sigma_t}[u_i(s)]$$

For any $i \in N$ and any $\delta : S_i \to S_i$

$$\frac{1}{T}\left[\sum_{t=1}^{T}\sum_{a\in A} p_i^t(a)\pi_t(\delta(a)) - \sum_{t=1}^{T}\sum_{a\in A} p_i^t(a)\,\pi_t(a)\right] \leq \varepsilon$$

$$\Rightarrow \frac{1}{T}\left[\sum_{t=1}^{T} \mathop{\mathbb{E}}_{s\sim\sigma_t}\left[u_i\left(\delta(s_i), s_{-i}\right)\right] - \sum_{t=1}^{T} \mathop{\mathbb{E}}_{s\sim\sigma_t}\left[u_i(s)\right]\right] \leq \varepsilon$$

$$\Rightarrow \mathop{\mathbb{E}}_{s\sim\sigma}\left[u_i\left(\delta(s_i), s_{-i}\right)\right] - \mathop{\mathbb{E}}_{s\sim\sigma}\left[u_i(s)\right] \leq \varepsilon$$

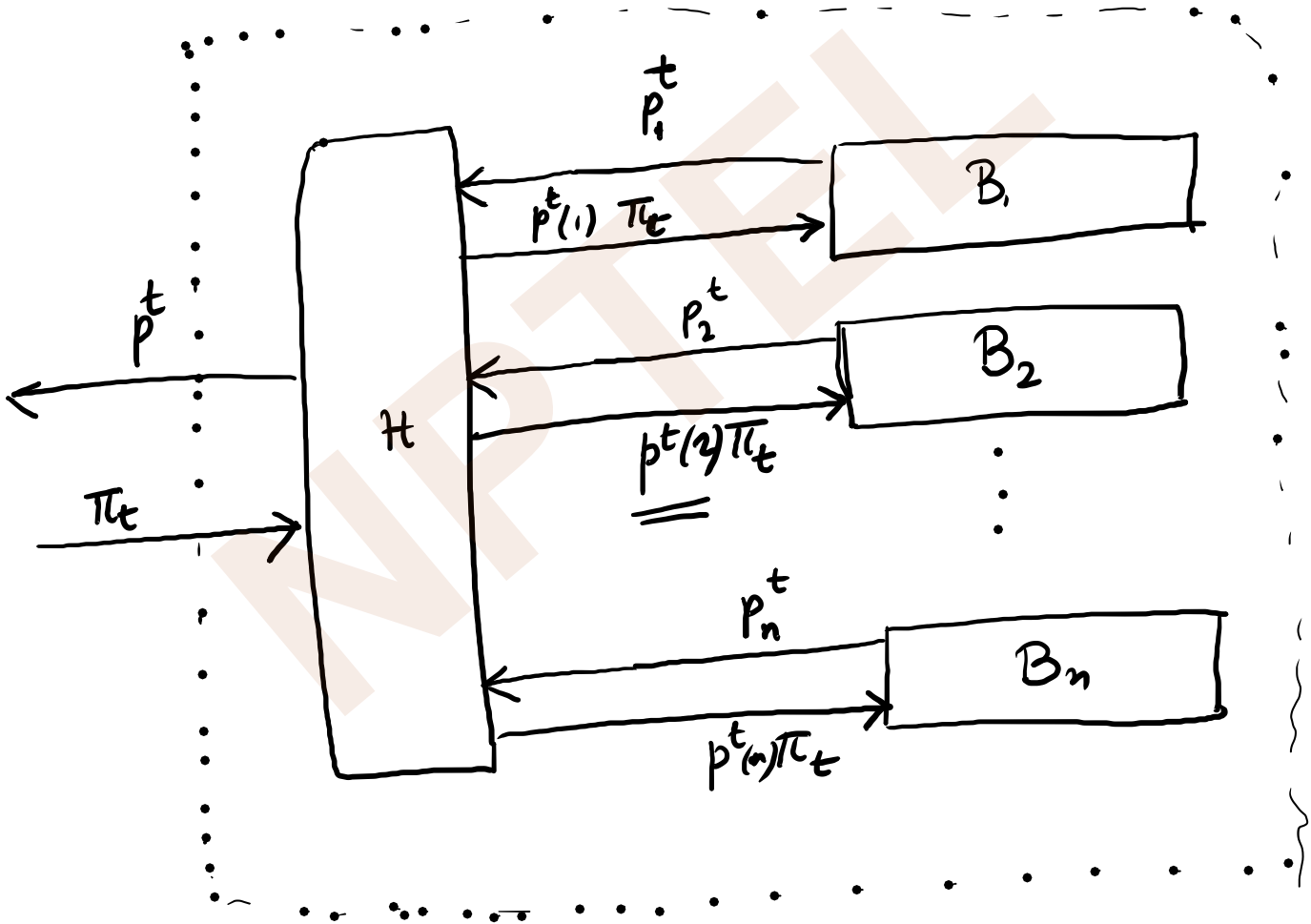
Q. Does there exist a no-swap-regret algorithm?

Black box reduction from no-swap regret to no external regret:

Theorem: Let $|A| = n$. Suppose there exists an algorithm with time averaged external regret $R(T, n)$. Then there also exists an algorithm with time averaged swap regret $n\,R(T, n)$. In particular, if there exists a

no external regret algorithm (i.e. $\lim_{T\to\infty} R(T,n) = 0$), then there also exists a no swap regret algorithm.

Proof: Let the action set of the player (of the no / swap regret algorithm) be $A = \{1, \ldots, n\}$. Let B be an algorithm with time-averaged external regret $R(T,n)$. Take n copies of B; let us call them $B_1, \ldots, B_n$. We will build a "master algorithm" H using $B_1, \ldots, B_t$


p_1^t
B_1
$p^t(1)\ \pi_t$
p^t
p_2^t
B_2
H
$p^t(2)\pi_t$
π_t
p_n^t
B_n
$p^t(n)\pi_t$

The time-averaged expected pay off of the master algorithm is

$$\frac{1}{T} \sum_{t=1}^{T} \sum_{i \in A} p^t(i) \cdot \pi_t(i)$$

Let $\delta : A \to A$ be any switching function. Then the time-averaged expected pay off of the master algorithm (modified by δ) is

$$\frac{1}{T} \sum_{t=1}^{T} \sum_{i \in A} p^t(i)\, \pi_t(\delta(i))$$

We need to show,

$$\frac{1}{T}\left[\sum_{t=1}^{T}\sum_{i\in A} p^t(i)\,\pi_t(\delta(i)) - \sum_{t=1}^{T}\sum_{i\in A} p^t(i)\,\pi_t(i)\right] \leq n \cdot R(T,n)$$

Since the external regret of each B_i is $R(T,n)$, we have the following

$$\frac{1}{T}\left(\sum_{t=1}^{T} p^t(i)\,\pi_t(\lambda) - \sum_{t=1}^{T}\sum_{j\in A} p_i^t(j)\,p^t(i)\,\pi_t(j)\right) \leq R(T,n)$$

$$\forall\ i \in A,\ \lambda \in A$$

Put $\lambda = \delta(i)$ in the above inequality.

$$\frac{1}{T}\left(\sum_{t=1}^{T} p^t(i)\, \pi_t(\delta(i)) - \sum_{t=1}^{T} \sum_{j \in A} p_i^t(j)\, p^t(i)\, \pi_t(j) \right) \leq R(T,n) \quad \forall i \in A.$$

Adding all the above n inequalities,

$$\frac{1}{T}\left(\sum_{t=1}^{T} \sum_{i \in A} p^t(i).\pi_t(\delta(i)) - \sum_{t=1}^{T} \sum_{i \in A} \sum_{j \in A} p_i^t(j)\, p^t(i)\, \pi_t(j) \right) \leq n R(T,n)$$

$$\Rightarrow \frac{1}{T}\left(\sum_{t=1}^{T} \sum_{i \in A} p^t(i)\, \pi_t(\delta(i)) - \sum_{t=1}^{T} \sum_{i \in A} \pi_t(i) \sum_{j \in A} p_j^t(i)\, p^t(j) \right) \leq n R(T,n)$$

Set p^t as a solution to the following system of linear equations.

$$p^t(i) = \sum_{j \in A} p^t_j(i)\, p^t(j) \qquad \forall\ i \in A$$

$p^t_j(i)$ — like constants.

$$\sum_{i \in A} p^t(i) = 1$$

Corollary: $|A| = n$. There exists an algorithm with swap regret $O\left(n \cdot \sqrt{\frac{\ln n}{T}}\right)$.

Lecture 7.4

Price of Anarchy

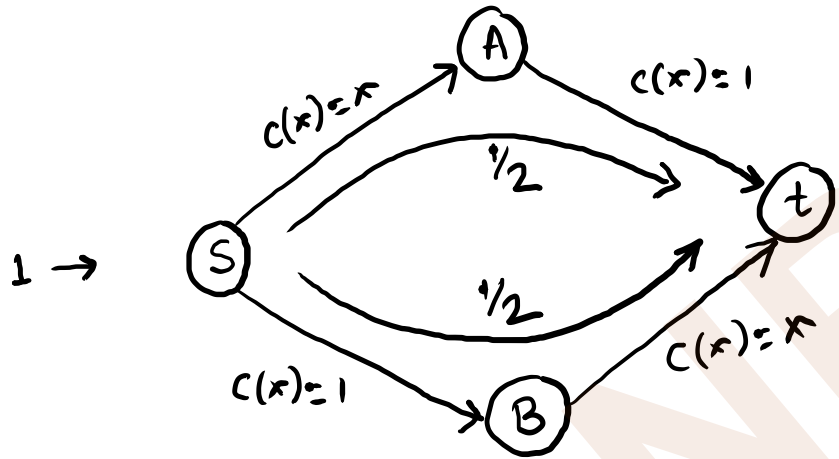
$$\Gamma = \langle N, (S_i)_{i \in N}, (u_i)_{i \in N} \rangle$$

Social welfare function: $w : \underset{i \in N}{X} S_i \longrightarrow \mathbb{R}_{\geq 0}$

Definition (PoA):

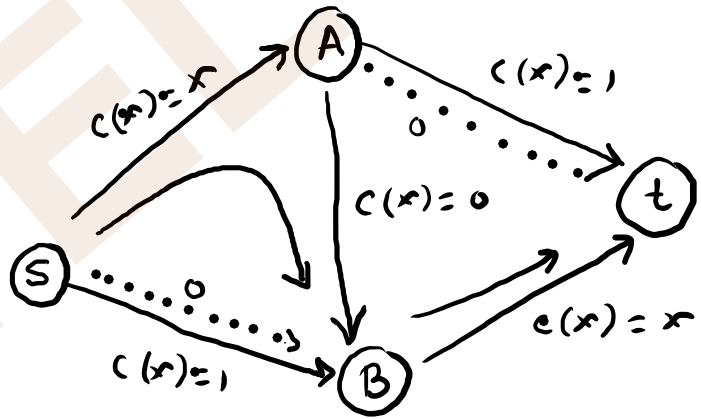
$$PoA = \frac{\max_{s \in \underset{i \in N}{X} S_i} w(s)}{\max_{s^{eq} \in PSNE(\Gamma)} w(s^{eq})}$$

Braess's Paradox



$\left(S \to A \to t : \frac{1}{2},\ S \to B \to t : \frac{1}{2}\right)$ is a PSNE

Total delay $= \frac{3}{2}$



$S \to A \to B$ always take a cost of x (≤ 1) which is at most the cost of $S \to B$

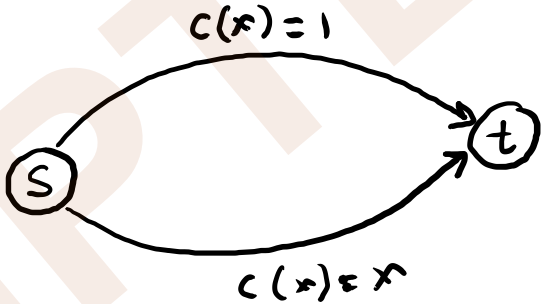
$S \to A \to B \to t$ is a weakly dominant strategy. Thus $(S \to A \to B \to t : 1)$ is a WDSE.

Total delay $= 2$

If the social welfare function is $\frac{1}{\text{total delay}}$, then

PoA is $\geq \frac{2}{3/2} = \frac{4}{3}$

Pigou's Network:



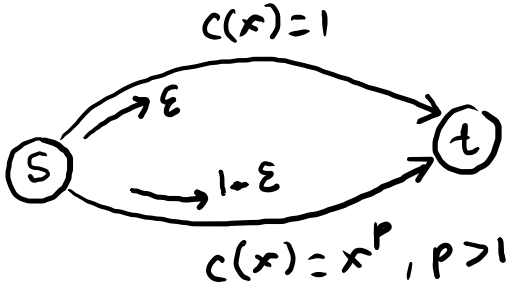
The bottom edge weakly dominates the top edge. So all traffic using the bottom edge is a WDSE.

Total cost = 1

$\frac{1}{2}$ traffic using the top edge and the other $\frac{1}{2}$ traffic

using the bottom edge incurs a cost of $\frac{1}{2} \cdot 1 + \frac{1}{2} \cdot \frac{1}{2} = \frac{3}{4}$.

$PoA \geq \frac{4}{3}$



All traffic following the bottom path is a WDSE. Total cost is 1.

Total cost is $\varepsilon \cdot 1 + (1-\varepsilon) \cdot (1-\varepsilon)^p$, $\varepsilon > 0$

$= \varepsilon + (1-\varepsilon)^{p+1}$

$$PoA \geq \frac{1}{\varepsilon + (1-\varepsilon)^{p+1}} \longrightarrow \infty \text{ as } p \to \infty$$

As non-linearity in the cost function increases, the PoA also increases.

PoA for Selfish Networks

Pigou's Network:

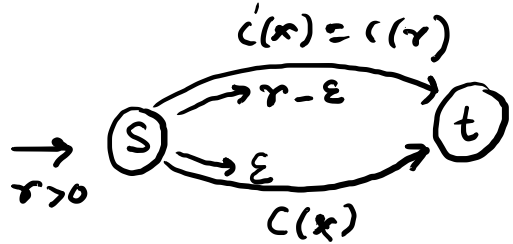
Define:

- It has two vertices: source s and sink t
- ——— edges from s to t.
- Traffic rate is r (>0)
- The cost of one edge is $c(x)$; $c(\cdot)$ is a non-negative, non-decreasing, continuous functio

- The cost of the other edge is $c(r)$.

PoA of Pigou's network:

$$\alpha(C) = \sup_{\varepsilon \in [0,r]} \left[\frac{r\, c(r)}{\varepsilon\, c(\varepsilon) + (r-\varepsilon)\, c(r)} \right]$$



Observe that, $\varepsilon\, c(\varepsilon) + (r-\varepsilon)\, c(r)$ is non-decreasing in $\varepsilon \in [r, \infty)$ since $c(\cdot)$ is a non-decreasing function.

$$\alpha(c) = \sup_{\varepsilon \geq 0} \frac{r\, c(r)}{\varepsilon\, c(\varepsilon) + (r-\varepsilon)\, c(r)}$$

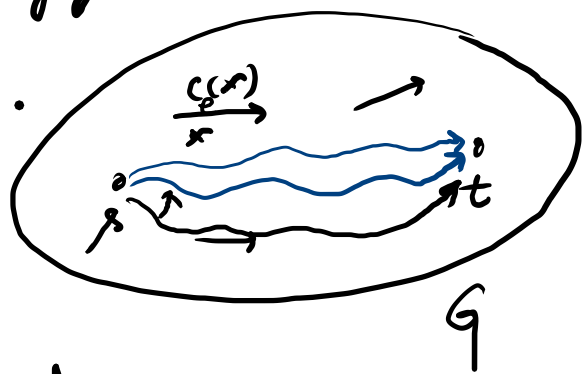
We let the cost function c belong to a class of functions $\mathcal{C}$.

$$\alpha(\mathcal{C}) = \sup_{c \in \mathcal{C}} \sup_{r>0} \alpha(c)$$

$$= \sup_{c \in \mathcal{C}} \sup_{r>0} \sup_{\varepsilon \geq 0} \left[\frac{r\,c(r)}{\varepsilon c(\varepsilon) + (r-\varepsilon)c(r)} \right]$$

Theorem: For any class $\mathcal{C}$ of cost functions, the PoA of any network with one source and one destination where all the cost functions belong to $\mathcal{C}$, is at most $\alpha(\mathcal{C})$.

Proof: Let G be a graph, $f : E[G] \to \mathbb{R}_{\geq 0}$ be an s-t flow in G. Observe that every strategy profile corresponds to a flow in G.



Observe that a flow f corresponds to a PSNE if and only if only shortest paths carry any flow.

$$f_{\bar{P}} > 0 \Rightarrow \bar{P} \in \underset{P \in \mathcal{P}}{\operatorname{argmin}} \left\{ \sum_{e \in P} c_e(f_e) \right\}$$

Let f^* be a flow which minimizes the sum of cost of all the edges.

$$f^* \in \underset{f \text{ is an } s\text{-}t \text{ flow of value } r}{\operatorname{argmin}} \left\{ \sum_{e \in E} f_e \, c_e(f_e) \right\}$$

Since PSNE flow f^{eq} only uses shortest paths according to the cost function $c_e(f_e^{eq})$, $e \in E[G]$, all the s-t path carrying any flow in f^{eq} must have the same cost.

$$\sum_{p \in P} f_p^{eq} \; c(f_p^{eq}) = rL \quad \text{, where } L \text{ is the cost of the min-cost path.}$$

—(1)

$$\sum_{p \in P} f_p^{*} \; \underbrace{c(f_p^{eq})}_{\geq L} \geq L \sum_{p \in P} f_p^{*} = rL \quad —(2)$$

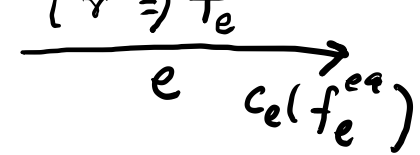
(1) − (2)

$$\sum_{p \in P} (f_p^{eq} - f_p^{*}) \, c(f_p^{eq}) \leq 0 \quad —(3)$$

$\alpha(C)$ is supremum over all $e \in C$, $r > 0$, $\varepsilon \geq 0$, we have the following for any edge $e \in E[G]$,

$$\alpha(C) \geq \frac{f_e^{eq} \cdot c_e(f_e^{eq})}{f_e^* c_e(f_e^*) + (f_e^{eq} - f_e^*) c_e(f_e^{eq})}$$

$$\Rightarrow f_e^* c_e(f_e^*) \geq \frac{f_e^{eq} \cdot c_e(f_e^{eq})}{\alpha(C)} - (f_e^{eq} - f_e^*) c_e(f_e^{eq})$$



$\varepsilon = f_e^*$

$r = f_e^{eq}$

$$\Rightarrow \underbrace{\sum_{e \in E} f_e^* c_e(f_e^*)}_{} \geq \frac{1}{\alpha(C)} \underbrace{\sum_{e \in E} f_e^{eq} c_e(f_e^{eq})}_{} - \underbrace{\sum_{e \in E} (f_e^{eq} - f_e^*) c_e(f_e^{eq})}_{\leq 0}$$

$$\Rightarrow \quad C(f^*) \geq \frac{C(f^{eq})}{\alpha(C)}$$

$$\Rightarrow \quad \frac{C(f^{eq})}{C(f^*)} \leq \alpha(C)$$

$$\Rightarrow \quad PoA \leq \alpha(C)$$