Step 2

$$\begin{aligned} \mathrm{E}_S[\Phi(S)] &= \mathrm{E}_S\left[\sup_{f\in\mathcal{F}}\left(\mathrm{E}[f] - \hat{\mathrm{E}}_S[f]\right)\right] \\ &\leq \mathrm{E}_{S,S'}\left[\sup_{f\in\mathcal{F}}\left(\hat{\mathrm{E}}_{S'}[f] - \hat{\mathrm{E}}_S[f]\right)\right] \end{aligned}$$

Proof:

- $\mathrm{E}_{S'}\left[\hat{\mathrm{E}}_{S'}[f]\right] = \mathrm{E}[f]$
- $\mathrm{E}_{S'}\left[\hat{\mathrm{E}}_{S}[f]\right] = \hat{\mathrm{E}}_S[f]$
- so:

$$\begin{aligned} \mathrm{E}_S\left[\sup_{f\in\mathcal{F}}\left(\mathrm{E}[f] - \hat{\mathrm{E}}_S[f]\right)\right] &= \mathrm{E}_S\left[\sup_{f\in\mathcal{F}} \mathrm{E}_{S'}\left[\hat{\mathrm{E}}_{S'}[f] - \hat{\mathrm{E}}_S[f]\right]\right] \\ &\leq \mathrm{E}_S\left[\mathrm{E}_{S'}\left[\sup_{f\in\mathcal{F}}\left(\hat{\mathrm{E}}_{S'}[f] - \hat{\mathrm{E}}_S[f]\right)\right]\right] \end{aligned}$$

- since: $\mathrm{E}[\max\{X_1,\ldots,X_n\}] \geq \max\{\mathrm{E}[X_1],\ldots,\mathrm{E}[X_n]\}$

Step 4

$$\mathrm{E}_{S,S',\boldsymbol{\sigma}}\left[\sup_{f\in\mathcal{F}}\left(\frac{1}{m}\sum_{i=1}^{m}\sigma_i\left(f(z_i')-f(z_i)\right)\right)\right]\leq 2\mathcal{R}_m(\mathcal{F})$$

Proof:

$$\begin{aligned}
&\mathrm{E}_{S,S',\boldsymbol{\sigma}}\left[\sup_{f\in\mathcal{F}}\left(\frac{1}{m}\sum_{i=1}^{m}\sigma_i\left(f(z_i')-f(z_i)\right)\right)\right]\\
&= \mathrm{E}_{S,S',\boldsymbol{\sigma}}\left[\sup_{f\in\mathcal{F}}\left(\frac{1}{m}\sum_{i=1}^{m}\sigma_i f(z_i')+\frac{1}{m}\sum_{i=1}^{m}(-\sigma_i)f(z_i)\right)\right]\\
&\leq \mathrm{E}_{S,S',\boldsymbol{\sigma}}\left[\sup_{f\in\mathcal{F}}\frac{1}{m}\sum_{i=1}^{m}\sigma_i f(z_i')+\sup_{f\in\mathcal{F}}\frac{1}{m}\sum_{i=1}^{m}(-\sigma_i)f(z_i)\right]\\
&= \underbrace{\mathrm{E}_{\cancel{S},S',\boldsymbol{\sigma}}\left[\sup_{f\in\mathcal{F}}\frac{1}{m}\sum_{i=1}^{m}\sigma_i f(z_i')\right]}_{=\mathcal{R}_m(\mathcal{F})}+\underbrace{\mathrm{E}_{S,\cancel{S'},\boldsymbol{\sigma}}\left[\sup_{f\in\mathcal{F}}\frac{1}{m}\sum_{i=1}^{m}(-\sigma_i)f(z_i)\right]}_{=\mathcal{R}_m(\mathcal{F})}
\end{aligned}$$