# SaddlePoint Form of the InverseRL Objective

**Theorem 1.** *Let*

$$J(\pi) = \ D_f(\mu_\pi \| \mu_E) \ - \ \lambda\, \mathcal{H}(\pi),$$

*where* $\mathcal{H}(\pi) = -\mathbb{E}_{\mu_\pi}[\log \pi]$. *Then*

$$\min_\pi J(\pi) \ = \ - \ \max_r \ \min_\pi \Big\{ \mathbb{E}_{\mu_E}\big[f^*(-r)\big] \ + \ \mathbb{E}_{\mu_\pi}\big[r - \lambda \log \pi\big] \Big\}.$$

*Proof.* **1. Variational representation of the** $f$**divergence.**
By Fenchel–Young,

$$f(t) = \sup_u \{u\, t - f^*(u)\},$$

so with $p = \frac{d\mu_\pi}{d\mu_E}$,

$$D_f(\mu_\pi \| \mu_E) = \mathbb{E}_{\mu_E}[f(p)] = \sup_g \Big\{ \mathbb{E}_{\mu_E}[g\, p] - \mathbb{E}_{\mu_E}[f^*(g)] \Big\} = \sup_g \Big\{ \mathbb{E}_{\mu_\pi}[g] - \mathbb{E}_{\mu_E}[f^*(g)] \Big\}.$$

Hence

$$J(\pi) = \sup_g \Big\{ \mathbb{E}_{\mu_\pi}[g] - \mathbb{E}_{\mu_E}[f^*(g)] \Big\} \ + \ \lambda\, \mathbb{E}_{\mu_\pi}[\log \pi].$$

**2. Change of variable** $g = -r$**.**
Set $g(s,a) = -r(s,a)$. Then

$$\mathbb{E}_{\mu_\pi}[g] = -\mathbb{E}_{\mu_\pi}[r], \quad -\mathbb{E}_{\mu_E}[f^*(g)] = -\mathbb{E}_{\mu_E}[f^*(-r)],$$

so

$$J(\pi) = \sup_r \Big\{ -\mathbb{E}_{\mu_E}[f^*(-r)] - \mathbb{E}_{\mu_\pi}[r] + \lambda\, \mathbb{E}_{\mu_\pi}[\log \pi] \Big\} = \sup_r \Big\{ -\mathbb{E}_{\mu_E}[f^*(-r)] - \mathbb{E}_{\mu_\pi}[\, r - \lambda \log \pi\,] \Big\}.$$

**3. Swap** $\min_\pi$ **and** $\sup_r$**.**
The inner function

$$L(\pi, r) = -\mathbb{E}_{\mu_E}[f^*(-r)] - \mathbb{E}_{\mu_\pi}[\, r - \lambda \log \pi\,]$$

is convex in $\pi$ and concave in $r$. By Sions minimax theorem,

$$\min_{\pi} \sup_{r} L(\pi, r) = \sup_{r} \min_{\pi} L(\pi, r).$$

Hence

$$\min_{\pi} J(\pi) = \sup_{r} \min_{\pi} \Big\{ -\mathbb{E}_{\mu_E}[f^*(-r)] - \mathbb{E}_{\mu_\pi}[\, r - \lambda \log \pi \,] \Big\}.$$

**4. Rearranging constants.**
Since $-\mathbb{E}_{\mu_E}[f^*(-r)]$ is independent of $\pi$,

$$\min_{\pi} J(\pi) = \sup_{r} \Big\{ - \mathbb{E}_{\mu_E}[f^*(-r)] - \max_{\pi} \mathbb{E}_{\mu_\pi}[\, r - \lambda \log \pi \,] \Big\} = - \inf_{r} \Big\{ \mathbb{E}_{\mu_E}[f^*(-r)] + \max_{\pi} \mathbb{E}_{\mu_\pi}[\, r - \lambda \log$$

which is algebraically equivalent to the claimed saddlepoint form. $\square$