Problem 2: Online optimization in continuous time. Consider the following online learning problem: At every instance $t \geq 0$ **(continuous time)**, a learner selects an action x_t from a compact, convex subset $\mathcal{X}$ of $\mathbb{R}^d$. The agent's loss at time t is given by $f_t(x_t)$, where each loss function $f_t: \mathcal{X} \to \mathbb{R}$, $t \geq 0$ is assumed convex; throughout this problem, we also assume that the map $(t, x) \mapsto f_t(x)$ is jointly continuous and differentiable with respect to x.

As usual, the agent's regret at time T is defined as

$$\mathrm{Reg}(T) = \max_{x \in \mathcal{X}} \int_0^T [f_s(x_s) - f_s(x)]\, ds,$$

and our aim will be to design an online learning policy that leads to **no regret**, i.e., such that $\mathrm{Reg}(T) = o(T)$ against any stream of loss functions f_t, $t \geq 0$.

We will focus on a continuous-time version of the **dual averaging** policy we examined in class. To begin, we say that a strictly convex function $\phi: \mathbb{R}^d \to \mathbb{R}$ is a **link function** for $\mathcal{X}$ if it is differentiable and

- $\nabla\phi(y) \in \mathcal{X}$ for all $y \in \mathbb{R}^d$.
- For every $x \in \operatorname{ri}\mathcal{X}$, there exists $y \in \mathbb{R}^d$ such that $x = \nabla\phi(y)$.

[NB: In the above and in what follows, $\operatorname{ri}\mathcal{X}$ denotes the relative interior of $\mathcal{X}$.]

(Q1) Show that $\phi(y) = \log \sum_{j=1}^d \exp(y_j)$ is a link function for the simplex $\mathcal{X} = \{x \in \mathbb{R}_+^d : \sum_{j=1}^d x_j = 1\}$. **[5 points]**

(Q2) Using the above result as a basis (or otherwise), find a link function for the cube $\mathcal{X} = [0,1]^d$. **[5 points]**

(Q3) Define the **convex conjugate** of ϕ as

$$h(x) = \sup_{y \in \mathbb{R}^d} \{\langle y, x\rangle - \phi(y)\}, \quad x \in \operatorname{ri}\mathcal{X}.$$

Show that the supremum above is attained and conclude that $h(x) < \infty$ for all $x \in \operatorname{ri}\mathcal{X}$. **[5 points]**

To continue, we will make the following blanket assumption:

- h is C^1-smooth and $\sup_{x \in \operatorname{ri}\mathcal{X}} h(x) < \infty$.

(Q4) Calculate the conjugate of $\phi(y) = \log \sum_{j=1}^d \exp(y_j)$ and show that the above assumption is satisfied. **[5 points]**

(Q5) Similarly for the mirror function you provided in Question (2), i.e., when $\mathcal{X} = [0,1]^d$. **[5 points]**

With all this at hand, the **dual averaging** policy is defined as

$$\begin{aligned} y_t &= -\int_0^t \nabla f_s(x_s)\, ds, \\ x_t &= Q(y_t), \end{aligned} \tag{DA}$$

where $Q = \nabla\phi$. To analyze this policy, introduce the so-called **Fenchel coupling**

$$F_p(t) = h(p) + \phi(y_t) - \langle y_t, p\rangle, \quad p \in \operatorname{ri}\mathcal{X}, t \geq 0.$$

(Q6) Show that

$$\frac{dF_p}{dt} = \langle \nabla f_t(x_t), p - x_t\rangle$$

[5 points]

(Q7) Conclude that

$$\mathrm{Reg}(T) \leq \phi(0) + \sup h$$

i.e., (DA) leads to **constant regret.** **[5 points]**

(Q8) Compute this bound when $\phi(y) = \log \sum_{j=1}^d \exp(y_j)$ and compare it to the corresponding bound for the exponential weights algorithm in discrete time. What is the main difference? Comment briefly on what you think is the source of this difference. **[5 points]**