1. **(3 marks)** For any $w, z$ show that $\langle w, z \rangle \leq ||w||_1 ||z||_1$ where $||.||_1$ is the dual norm of $||.||$. 

2. **(3 marks)** Consider an algorithm that enjoys a regret bound of the form $\alpha \sqrt{T}$, but its parameters require the knowledge of $T$. The doubling trick, described below, enables us to convert such an algorithm into an algorithm that does not need to know the time horizon. The idea is to divide the time into periods of increasing size and run the original algorithm on each period.

**Algorithm 1 Doubling Trick**

**Input:** An algorithm $A$ whose parameter depend on $n$
for $m = 0, 1, 2, ...$

Run $A$ for $2^m$ rounds for $t = 2^m, 2^{m+1}, ... 2^{m+1} - 1$.

Show that using the Doubling trick algorithm the total regret obtained if you stop after $m$ rounds is $\frac{\sqrt{2^m}}{\sqrt{2} - 1} \alpha \sqrt{m}$.

3. **(4 marks)** Consider the following algorithm known as Online Mirror Descent

**Algorithm 2 Online Mirror Decent**

**Input:** A link function $g : R^d \rightarrow S$

**Initialize:** $\theta_1 = 0$
for $t = 0, 1, 2, ..$

predict $w_t = g(\theta_t)$
update $\theta_{t+1} = \theta_t - z_t$ where $z_t \in \partial f_t(w_t)$

Show that when $g(\theta) = \eta \theta$, the above algorithm gives the same update rules as the online gradient descent.