Let $f : \mathbb{R}^d \to \mathbb{R}$ be a function. We assume that

1. f is convex ;
2. f has a global minimizer x_* ;
3. f is differentiable and, for any $x \in \mathbb{R}^d$,

$$||\nabla f(x)||_2 \leq 1.$$

We fix a starting point x_0 and run gradient descent from this point, with a sequence of positive stepsizes $(h_k)_{k\in\mathbb{N}}$:

$$x_{k+1} = x_k - h_k \nabla f(x_k).$$

1. a) Show that, for any $k \in \mathbb{N}$,

$$f(x_k) - f(x_*) \leq \langle \nabla f(x_k), x_k - x_* \rangle .$$

b) Show that, for any $k \in \mathbb{N}$,

$$||x_{k+1} - x_*||_2^2 \leq ||x_k - x_*||_2^2 - 2h_k(f(x_k) - f(x_*)) + h_k^2 ||\nabla f(x_k)||_2^2.$$

c) Show that, for any $n \in \mathbb{N}$,

$$2\sum_{k=0}^{n} h_k(f(x_k) - f(x_*)) \leq ||x_0 - x_*||_2^2 - ||x_{n+1} - x_*||_2^2 + \sum_{k=0}^{n} h_k^2 ||\nabla f(x_k)||_2^2.$$

d) For any n, let $k_n \in \{0, \dots, n\}$ be such that

$$f(x_{k_n}) = \min_{k=0,\dots,n} f(x_k).$$

Show that, for any n,

$$2(f(x_{k_n}) - f(x_*)) \left(\sum_{k=0}^{n} h_k \right) \leq ||x_0 - x_*||_2^2 - ||x_{n+1} - x_*||_2^2 + \sum_{k=0}^{n} h_k^2 ||\nabla f(x_k)||_2^2.$$

e) Show that, for any n,

$$2(f(x_{k_n}) - f(x_*)) \left(\sum_{k=0}^{n} h_k \right) \leq ||x_0 - x_*||_2^2 + \sum_{k=0}^{n} h_k^2.$$

2. In this question, we assume that, for any k, $h_k = \frac{1}{\sqrt{k+1}}$. Show that, for any n,

$$f(x_{k_n}) - f(x_*) \leq \frac{||x_0 - x_*||_2^2 + 2 + \log(n)}{\sqrt{n+2}}.$$

Hint : You can use the fact that, for any n,

$$\sum_{k=1}^{n+1} \frac{1}{k} \leq 2 + \log(n) \quad \text{and} \quad \sum_{k=1}^{n+1} \frac{1}{\sqrt{k}} \geq \frac{\sqrt{n+2}}{2}.$$

3. In this question, we assume that the sequence of stepsizes is constant : there exists $\eta > 0$ such that, for any $k \in \mathbb{N}$, $h_k = \eta$.

Give an example of a function f satisfying properties 1, 2, 3, and a starting point x_0 such that

$$f(x_{k_n}) - f(x_*) \overset{n\to+\infty}{\not\to} 0.$$

Hint : Define

$$\begin{aligned} f : x \in \mathbb{R} \quad \to \quad & |x| - \frac{\epsilon}{2} && \text{if } |x| \geq \epsilon; \\ & \frac{x^2}{2\epsilon} && \text{if } |x| \leq \epsilon, \end{aligned}$$

for some $\epsilon > 0$ small enough.