Name: **Present group members:**

Some useful facts about Lipschitz functions:

- The Lipshitz requirement can be reframed as $\frac{|f(\mathbf{x})-f(\mathbf{y})|}{\|\mathbf{x}-\mathbf{y}\|} \leq L$, and that thing on the left is related to the derivative. That means that if the derivative is bounded, then the function is Lipschitz with constant L.
- If $f : U \subset \mathbb{R}^n \to \mathbb{R}$, and $g : [a, b] \subset \mathbb{R} \to \mathbb{R}$ are Lipschitz continuous with constants L_1, and L_2, respectively, then their composition $g \circ f = g(f)$ is Lipschitz continuous with constant L_1L_2.
- Lipschitz functions are closed under common algebraic operations: addition, subtraction, and scalar multiplication. This means, for example, if you add two Lipshitz functions then the resulting function from these operations is also Lipschitz.

Worksheet 4-3: Q1

For each of the following state (with justification):

(a) If the function is Lipschitz (and if so what is L)

(b) If the derivative is Lipschitz (and if so, what is L),

(c) If the function is in $C_L^{1,1}$ (and if so, what is L).

1. $f(x) = mx + b$ (linear function)
   - The derivative is $f'(x) = m$, so the function f is Lipschitz with constant $|m|$
   - The second derivative is $f''(x) = 0$, which is Lipschitz with constant 0.
   - Because the derivative (AKA gradient) has a bounded derivative (here 0), f is in $C_0^{1,1}$.

2. $f(x) = \sqrt{x}$
   - $f'(x) = \frac{1}{2\sqrt{x}}$ which is unbounded as $x \to 0$, so f is not Lipschitz.
   - $f''(x) = -\frac{1}{4x^{3/2}}$ which is also unbounded as $x \to 0$, so f is not in $C^{1,1}$.
   - f'' isn't Lipschitz, so f is not in $C^{1,1}$.

3. $f(x) = x^2$

- $f'(x) = 2x$ which is unbounded as $x \to \infty$, so f is not Lipschitz.
- $f''(x) = 2$ which is bounded, so f' Lipschitz with constant 0.
- Since f' is Lipschitz with constant 0, f is in $C_0^{1,1}$.

4. $f(x) = \sin(x)$

- Since $f'(x) = \cos(x)$ which is bounded by 1, f is Lipschitz with constant $L = 1$.
- The second derivative is $f''(x) = -\sin(x)$ which is also bounded by 1, so f' is Lipschitz with constant $L = 1$.
- $f \in C_1^{1,1}$ with constant $L = 1$.

5. $f(x) = \exp(-x)$ for $x \in [0, \infty)$

- $f'(x) = -\exp(-x)$. While $f'(x)$ is unbounded on all of $\mathbb{R}$, we're limited to $[0, \infty)$ where $f'(x)$ is bounded by 1. So, f is Lipschitz with constant $L = 1$.
- The second derivative is $f''(x) = \exp(-x)$ which is also bounded by 1 in $[0, \infty)$, so f' is Lipschitz with constant $L = 1$.
- $f \in C_1^{1,1}$ with constant $L = 1$.

6. $f(x) = 2\sin(x) - 10.9y^2 + \pi\exp(-(x^2 + y^2))$

- We already showed that $\sin(x)$ is Lipschitz with constant 1. The function $-10.9y^2$ is not Lipschitz because its derivative is unbounded as $y \to \infty$. The function $\pi\exp(-(x^2 + y^2))$ has bounded derivatives of all orders, so it is Lipschitz. However, since one of the terms is not Lipschitz, the whole function is not Lipschitz.
- For each of the pieces, we showed that the derivative is Lipshitz (since the second derivatives are bounded). Therefore, the whole function has Lipschitz derivative.
- Since the derivative is Lipschitz, the function is in $C_L^{1,1}$ for some L.

Worksheet 4-3: Q2

Given the function $f(x, y) = x^2 + y^4$, answer the following questions.

i) Write down the first 3 iterations of the gradient descent algorithm starting at $x_0 = (1, 1)$ using constant step size $t = \frac{1}{2}$.

- $x_{k+1} = x_k - t_k \nabla f(x_k)$
- $\nabla f(x, y) = \begin{bmatrix} 2x \\ 4y^3 \end{bmatrix}$
- Iteration 0: $x_0 = (1, 1)$
- Iteration 1:

$$\begin{aligned} x_1 &= x_0 - \frac{1}{2}\nabla f(x_0) \\ &= \begin{bmatrix} 1 \\ 1 \end{bmatrix} - \frac{1}{2}\begin{bmatrix} 2 \\ 4 \end{bmatrix} \\ &= \begin{bmatrix} 0 \\ -1 \end{bmatrix} \end{aligned}$$

- Iteration 2:

$$\begin{aligned} x_2 &= x_1 - \frac{1}{2}\nabla f(x_1) \\ &= \begin{bmatrix} 0 \\ -1 \end{bmatrix} - \frac{1}{2}\begin{bmatrix} 0 \\ -4 \end{bmatrix} \\ &= \begin{bmatrix} 0 \\ 1 \end{bmatrix} \end{aligned}$$

- Iteration 3:

$$\begin{aligned} x_3 &= x_2 - \frac{1}{2}\nabla f(x_2) \\ &= \begin{bmatrix} 0 \\ 1 \end{bmatrix} - \frac{1}{2}\begin{bmatrix} 0 \\ 4 \end{bmatrix} \\ &= \begin{bmatrix} 0 \\ -1 \end{bmatrix} \end{aligned}$$

ii) Comment on the results, are the iterations converging? *The iterations are oscillating between the points (0,1) and (0,-1). This means they will never converge to the actual minimum at (0,0).*

iii) Check whether the function satisfies the assumptions of the gradient method convergence theorem, and use your analysis to comment on the results of the iterations.

- Things to check:
- Is $f \in C_L^{1,1}$?
  - $\nabla f(x) = \begin{bmatrix} 2x \\ 4y^3 \end{bmatrix}$
  - $\nabla^2 f(x) = \begin{bmatrix} 2 & 0 \\ 0 & 12y^2 \end{bmatrix}$
  - $\|\nabla^2 f(x)\|$ is unbounded as $y \to \infty$. From the theorem on the slides, this means $f \notin C_L^{1,1}$ for any L.
- Is there an $m \in \mathbb{R}$ such that $f(x) > m$ for all $x \in \mathbb{R}^n$?
  - $f(x, y) = x^2 + y^4 \geq 0$ for all $(x, y) \in \mathbb{R}^2$, so we can take $m = 0$.
- However, since $f \notin C_L^{1,1}$, we cannot apply the gradient method convergence theorem to guarantee that the iterations will converge. This is consistent with what we observed in part (ii).