1. Equality of mixed partial derivatives

First a cautionary tale.

Example 1.1. *Let* $f(x_1, x_2) = \frac{x_1^2 - x_2^2}{x_1^2 + x_2^2}$. *Observe that*

$$\lim_{x_1 \to 0} \lim_{x_2 \to 0} f(x_1, x_2) = \lim_{x_1 \to 0} \frac{x_1^2}{x_1^2} = 1.$$

However,

$$\lim_{x_2 \to 0} \lim_{x_1 \to 0} f(x_1, x_2) = \lim_{x_1 \to 0} \frac{-x_2^2}{x_2^2} = -1.$$

The moral? One cannot generally switch the order in which one takes limits and expect to get the same answer.

My real aim here is to prove the following theorem which tells us that order is irrelevant when we take multiple partial derivatives of a 'decent' function of several variables.

Theorem 1.2. *Suppose that* $f : \mathbf{R}^n \to \mathbf{R}$ *is* C^2. *Then for any* $1 \le i, j \le n$ *and any* $a \in \mathbf{R}^n$ *in the domain of* f, *one has*

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(a) = \frac{\partial^2 f}{\partial x_j \partial x_i}(a).$$

My proof is quite similar to Shifrin's, but (in my humble opinion) mine ends better. In any case, the main thing is to show that one can reverse the order of the two limits involved in taking a second partial derivative.

Proof. To start with, note that since we are only considering derivatives of f with respect to x_i and x_j, we might as well assume that these are the *only* variables on which f depends. That is, it suffices to assume that $n = 2$ in the statement of the theorem, fix a point $a = (a_1, a_2)$ in the domain of f and show that

$$\frac{\partial^2 f}{\partial x_1 \partial x_2}(a) = \frac{\partial^2 f}{\partial x_2 \partial x_1}(a).$$

To this end, I go back to the definition of derivative, applying it to both partial derivatives:

$$\begin{aligned}
\frac{\partial^2 f}{\partial x_1 \partial x_2}(a) &= \lim_{h_1 \to 0} \frac{\frac{\partial f}{\partial x_1}(a_1 + h_1, a_2) - \frac{\partial f}{\partial x_1}(a_1, a_2)}{h_1} \\
&= \lim_{h_1 \to 0} \frac{\lim_{h_2 \to 0} \left(\frac{f(a_1+h_1, a_2+h_2) - f(a_1+h_1, a_2)}{h_2} \right) - \lim_{h_2 \to 0} \left(\frac{f(a_1, a_2+h_2) - f(a_1, a_2)}{h_2} \right)}{h_1} \\
&= \lim_{h_1 \to 0} \lim_{h_2 \to 0} \frac{f(a_1 + h_1, a_2 + h_2) - f(a_1 + h_1, a_2) - f(a_1, a_2 + h_2) + f(a_1, a_2)}{h_1 h_2}
\end{aligned}$$

Let me (for brevity's sake) call the quantity inside the last limit $Q(h_1, h_2)$.
Unnecessary motivational digression: *Similarly, when the partial derivatives are reversed, one finds:*

$$\frac{\partial^2 f}{\partial x_2 \partial x_1}(a) = \lim_{h_2 \to 0} \lim_{h_1 \to 0} Q(h_1, h_2)$$

That is, we get the same thing as before, except that the order of the limits is reversed. If we could switch the limits, we'd be home-free. But without justification, we can't. Instead we take a less direct but more justifiable approach that relies on the mean value theorem.

Lemma 1.3. *For each* $h = (h_1, h_2) \in \mathbf{R}^2$, *there exists* $\tilde{h} = (\tilde{h}_1, \tilde{h}_2)$ *inside the rectangle determined by* h *and* 0 *such that*

$$Q(h) = \frac{\partial^2 f}{\partial x_2 \partial x_1}(a + \tilde{h})$$

Proof. Note (i.e. really—check it!) that we can rewrite

$$Q(h_1, h_2) = \frac{1}{h_2} \frac{g(a_1 + h_1) - g(a_1)}{h_1}$$

where $g : \mathbf{R} \to \mathbf{R}$ is given by $g(t) := f(t, a_2 + h_2) - f(t, a_2)$. In particular g is a differentiable function of one variable with derivative given by $g'(t) = \frac{\partial f}{\partial x_1}(t, a_2 + h_2) - \frac{\partial f}{\partial x_1}(t, a_2)$. So I can apply the mean value theorem, obtaining a number $\tilde{h}_1$ between 0 and h_1 such that

$$Q(h_1, h_2) = \frac{1}{h_2}\left(\frac{g(a_1 + h_1) - g(a_1)}{h_1}\right) = \frac{1}{h_2} g'(a_1 + \tilde{h}_1) = \frac{1}{h_2}\left(\frac{\partial f}{\partial x_1}(a_1 + \tilde{h}_1, a_2 + h_2) - \frac{\partial f}{\partial x_1}(a_1 + \tilde{h}_1, a_2)\right).$$

Applying the Mean Value Theorem a second time, to this last expression, gives me a number $\tilde{h}_2$ between 0 and h_2 such that

$$Q(h_1, h_2) = \frac{\partial^2 f}{\partial x_2 \partial x_1}(a_1 + \tilde{h}_1, a_2 + \tilde{h}_2)$$

□

To finish the proof of the theorem, I will use the convenient notation $A \approx_\epsilon B$ to mean that $A, B \in \mathbf{R}$ satisfy $|A - B| < \epsilon$. Note that (by the triangle inequality) we have 'approximate transitivity'—i.e. $A \approx_{\epsilon_1} B$ and $B \approx_{\epsilon_2} C$ implies $A \approx_{\epsilon_1 + \epsilon_2} C$.

It will suffice to show that

$$\frac{\partial^2 f}{\partial x_1 \partial x_2}(a) \approx_\epsilon \frac{\partial^2 f}{\partial x_2 \partial x_1}(a)$$

for every $\epsilon > 0$. So let $\epsilon > 0$ be given. By continuity of second partial derivatives, there exists $\delta > 0$ such that $\|h\| < \delta$ implies that

$$\left| \frac{\partial^2 f}{\partial x_2 \partial x_1}(a + h) - \frac{\partial^2 f}{\partial x_2 \partial x_1}(a) \right| < \frac{1}{3}\epsilon.$$

Using the definition of limit twice and then the above lemma, I therefore obtain that when h_1 and then h_2 are small enough,

$$\frac{\partial^2 f}{\partial x_1 \partial x_2}(a) = \lim_{h_1 \to 0} \lim_{h_2 \to 0} Q(h_1, h_2) \approx_{\epsilon/3} \lim_{h_2 \to 0} Q(h_1, h_2) \approx_{\epsilon/3} Q(h_1, h_2) = \frac{\partial^2 f}{\partial x_2 \partial x_1}(a + \tilde{h}) \approx_{\epsilon/3} \frac{\partial^2 f}{\partial x_2 \partial x_1}(a).$$

In short,

$$\frac{\partial^2 f}{\partial x_1 \partial x_2}(a) \approx_\epsilon \frac{\partial^2 f}{\partial x_2 \partial x_1}(a),$$

which is what I sought to show. □